

NONA EDIÇÃO

INTRODUÇÃO à GENÉTICA

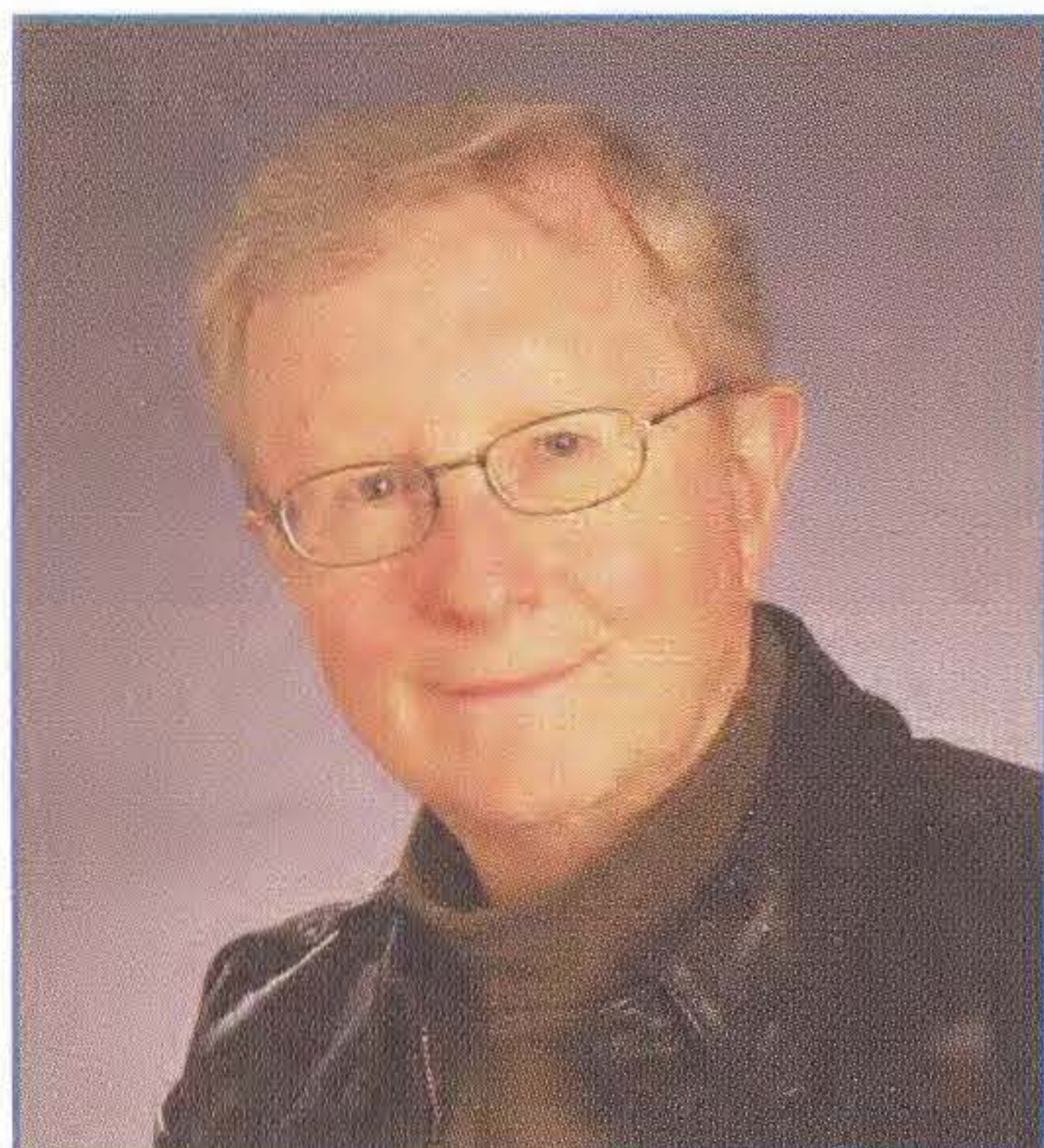
Anthony J. F. Griffiths • Susan R. Wessler • Richard C. Lewontin • Sean B. Carroll



GUANABARA  KOOGAN

Introdução à Genética

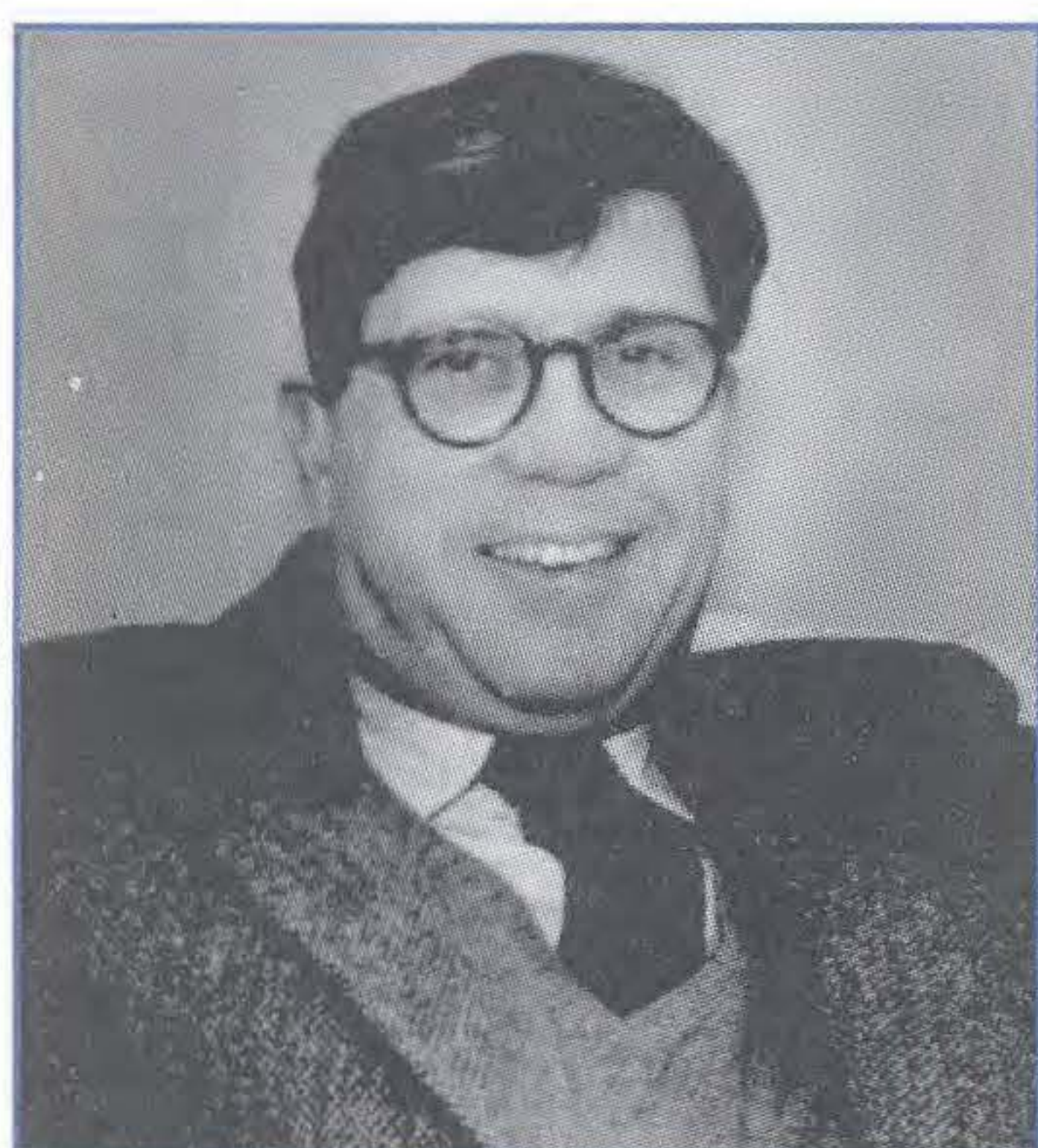
Sobre os Autores



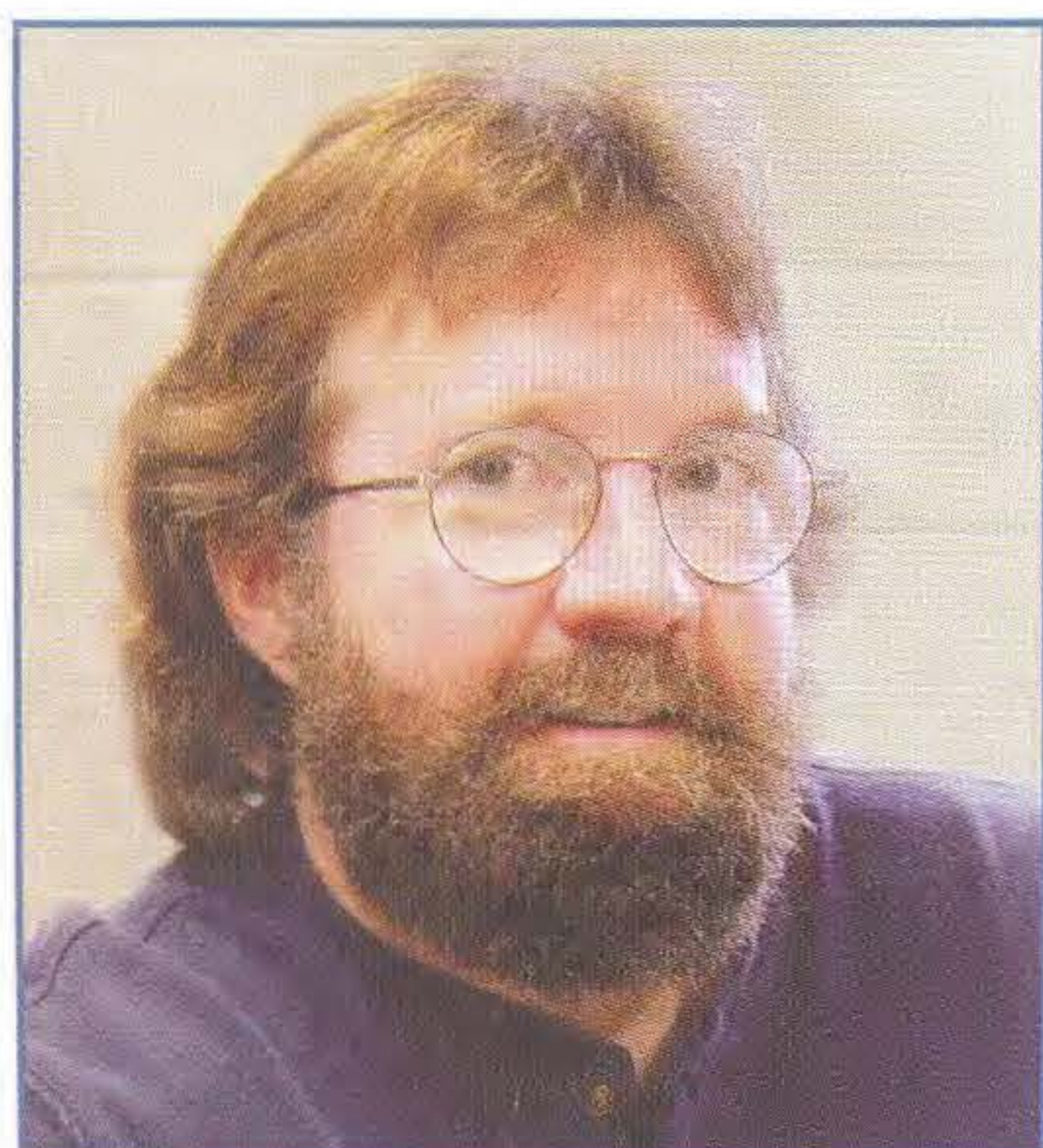
Anthony Griffiths é Professor Emérito da University of British Columbia, onde ensinou introdução à genética por 35 anos. Os desafios para ministrar esse curso levaram a um constante interesse em como os estudantes aprendem genética. Seu gosto pelas pesquisas está centrado na genética do desenvolvimento do fungo *Neurospora crassa*. Ele também gosta de lidar com a genética de populações das plantas locais. O Dr. Griffiths foi presidente da Genetics Society of Canada de 1987 a 1989, recebendo seu Award of Excellence em 1997. Recentemente serviu como secretário geral da International Genetics Federation.



Susan Wessler é Regents Professor of Plant Biology na University of Georgia, onde atua desde 1983. Leciona cursos de introdução à biologia e genética vegetal a estudantes de graduação e graduados. Seu interesse pelos métodos novos de ensino a levou a ser escolhida pelo Howard Hughes Medical Institute Professor em 2006. É co-autora do *The Mutants of Maize* (Cold Spring Harbor Laboratory Press) e de mais de 100 artigos baseados em pesquisas. Seu interesse científico enfoca os elementos de transposição e a estrutura e a evolução dos genomas. A Dra. Wessler foi eleita membro da National Academy of Sciences em 1998.



Richard Lewontin é Alexander Agassiz Research Professor na Harvard University. Tem ensinado genética, estatística e evolução na North Carolina State University, University of Rochester, University of Chicago e Harvard University. Sua principal área de pesquisa é a genética de populações e a genética evolutiva; introduziu métodos moleculares na genética de populações em 1966. Desde então, concentrou seus estudos na variação genética de proteínas e DNA dentro das espécies. O Dr. Lewontin foi presidente da Society for the Study of Evolution, American Society of Naturalists e Society for Molecular Biology and Evolution, e por alguns anos foi co-editor do *The American Naturalist*.



Sean Carroll é Professor of Molecular Biology and Genetics e Pesquisador no Howard Hughes Medical Institute na University of Wisconsin-Madison, onde ensina genética e biologia do desenvolvimento. A pesquisa do Dr. Carroll é centrada nos genes que controlam os padrões corpóreos e têm papéis importantes na evolução da diversidade animal. É autor de vários livros, inclusive o *The Making of the Fittest* (2006, W.W. Norton) e *Endless Forms Most Beautiful: The New Science of Evo Devo* (2005, W.W. Norton). Este último foi finalista do 2005 Los Angeles Times Book Prize (Science and Technology) e do 2006 National Academy of Sciences Communication Award. O Dr. Carroll também é co-autor com Jen Grenier e Scott Weatherbee do livro *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (2nd ed.; Blackwell Scientific) e autor ou co-autor de mais de 100 artigos baseados em pesquisa.

INTRODUÇÃO À GENÉTICA

Nona Edição

Anthony J. F. Griffiths

University of British Columbia

Susan R. Wessler

University of Georgia

Richard C. Lewontin

Harvard University

Sean B. Carroll

Howard Hughes Medical Institute

University of Wisconsin



GUANABARA  KOOGAN

Os autores e a editora empenharam-se para citar adequadamente e dar o devido crédito a todos os detentores dos direitos autorais de qualquer material utilizado neste livro, dispondo-se a possíveis acertos caso, inadvertidamente, a identificação de algum deles tenha sido omitida.

Cover Image: Nipam H. Patel <http://www.patellab.org>

First published in the United States

by

W.H. FREEMAN AND CO., New York and Basingstoke

Copyright © 2008 by W.H. Freeman and Co. All Rights Reserved

Publicado originalmente nos Estados Unidos

por

W.H. FREEMAN AND CO., New York and Basingstoke

Copyright © 2008 by W.H. Freeman and Co. Todos os Direitos Reservados.

Direitos exclusivos para a língua portuguesa

Copyright © 2009 by

EDITORA GUANABARA KOOGAN S.A.

Uma editora integrante do GEN | Grupo Editorial Nacional

Reservados todos os direitos. É proibida a duplicação ou reprodução deste volume, no todo ou em parte, sob quaisquer formas ou por quaisquer meios (eletrônico, mecânico, gravação, fotocópia, distribuição na internet ou outros), sem permissão expressa da Editora.

Travessa do Ouvidor, 11

Rio de Janeiro, RJ — CEP 20040-040

Tel.: 21-3970-9480

Fax: 21-2221-3202

gbk@editoraguanabara.com.br

www.editoraguanabara.com.br

Editoração Eletrônica:  ANTHARES

CIP-BRASIL. CATALOGAÇÃO NA FONTE
SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ

I48

Introdução à genética / Anthony J. F. Griffiths... [et al.] ; [traduzido por Paulo A. Motta]. - Rio de Janeiro : Guanabara Koogan, 2008.
il.

Tradução de: Introduction to genetic analysis, 9th ed

Apêndices

ISBN 978-85-277-1497-6

1. Genética. 2. Hereditariedade. I. Griffiths, Anthony J. F.

08-3055.

CDD: 576.5

CDU: 575

22.07.08

25.07.08

007820

Traduzido por

Paulo A. Motta

Ex-Professor Adjunto do Departamento de Genética da UFRJ e do
Instituto de Biologia da UFF

Conteúdo Resumido

1. O Enfoque Genético à Biologia, 1

PARTE I GENÉTICA DA TRANSMISSÃO

2. Herança Monogênica, 25

3. Distribuição Independente de Genes, 75

4. Mapeamento de Cromossomos Eucarióticos por Recombinação, 109

5. A Genética de Bactérias e seus Vírus, 151

6. Interação Gênica, 187

PARTE II DO DNA AO FENÓTIPO

7. DNA: Estrutura e Replicação, 225

8. RNA: Transcrição e Processamento, 253

9. Proteínas e sua Síntese, 275

10. Regulação da Expressão Gênica em Bactérias e seus Vírus, 301

11. Regulação da Expressão Gênica em Eucariontes, 329

12. O Controle Genético do Desenvolvimento, 355

13. Genomas e Genômica, 387

PARTE III MUTAÇÃO, VARIAÇÃO E EVOLUÇÃO

14. O Genoma Dinâmico: Elementos de Transposição, 417

15. Mutação, Reparo e Recombinação, 439

16. Alterações Cromossômicas em Larga Escala, 475

17. Genética de Populações, 517

18. Genética Quantitativa, 545

19. Genética Evolutiva, 579

PARTE IV TÉCNICAS

20. Isolamento e Manipulação de Gene, 609

Um Guia Resumido para Organismos-modelo, 647

Apêndice A: Nomenclatura Genética, 663

Apêndice B: Recursos de Bioinformática para Genética e Genômica, 665

Glossário, 667

Respostas de Problemas Seleccionados, 689

Índice Alfabético, 701

Conteúdo

1 O Enfoque Genético à Biologia, 1

1.1 Genética e as Perguntas da Biologia, 2

1.2 A Base Molecular da Informação Genética, 4

Especificando a sequência de aminoácidos de uma proteína, 5

Regulação gênica, 7

1.3 O Programa da Investigação Genética, 7

A necessidade da variação, 7

Começando com variação: Genética direta, 8

Começando com o DNA: Genética reversa, 11

1.4 Metodologias Usadas em Genética, 12

Visão geral, 12

Deteção de moléculas específicas de DNA, RNA e proteínas, 12

1.5 Organismos-modelo, 13

Lições dos primeiros organismos-modelo, 13

A necessidade de uma variedade de organismos-modelo, 14

1.6 Genes, o Ambiente e o Organismo, 17

Modelo I: Determinação genética, 17

Modelo II: Determinação ambiental, 17

Modelo III: Interação genótipo-ambiente, 18

O uso do genótipo e fenótipo, 18

Ruído de desenvolvimento, 19

Três níveis de desenvolvimento, 19



PARTE I GENÉTICA DA TRANSMISSÃO

2 Herança Monogênica, 25

2.1 Genes e Cromossomos, 27

2.2 Padrões de Herança Monogênica, 30

A lei de Mendel da segregação igual, 30

2.3 A Base Cromossômica dos Padrões de Herança Monogênica, 35

Herança monogênica em haplóides, 37

A base molecular da segregação e expressão monogênica, 38

2.4 Descobrindo Genes pela Observação das Proporções de Segregação, 47

Descobrindo um gene ativo no desenvolvimento da cor da flor, 48

Descobrindo um gene para desenvolvimento das asas, 48

Descobrindo um gene para a produção de esporos, 49

Os resultados da descoberta do gene, 49

Genética direta, 50

Prever as proporções da prole ou genótipos parentais aplicando os princípios da herança monogênica, 50

2.5 Padrões de Herança Monogênica Ligada ao Sexo, 50

Cromossomos sexuais, 51

Padrões de herança ligada ao sexo, 51

Boxe de Organismo-modelo – *Drosophila*, 52

Herança ligada ao X, 53

2.6 Análise de Heredogramas, 53

Distúrbios autossômicos recessivos, 55

Distúrbios autossômicos dominantes, 56

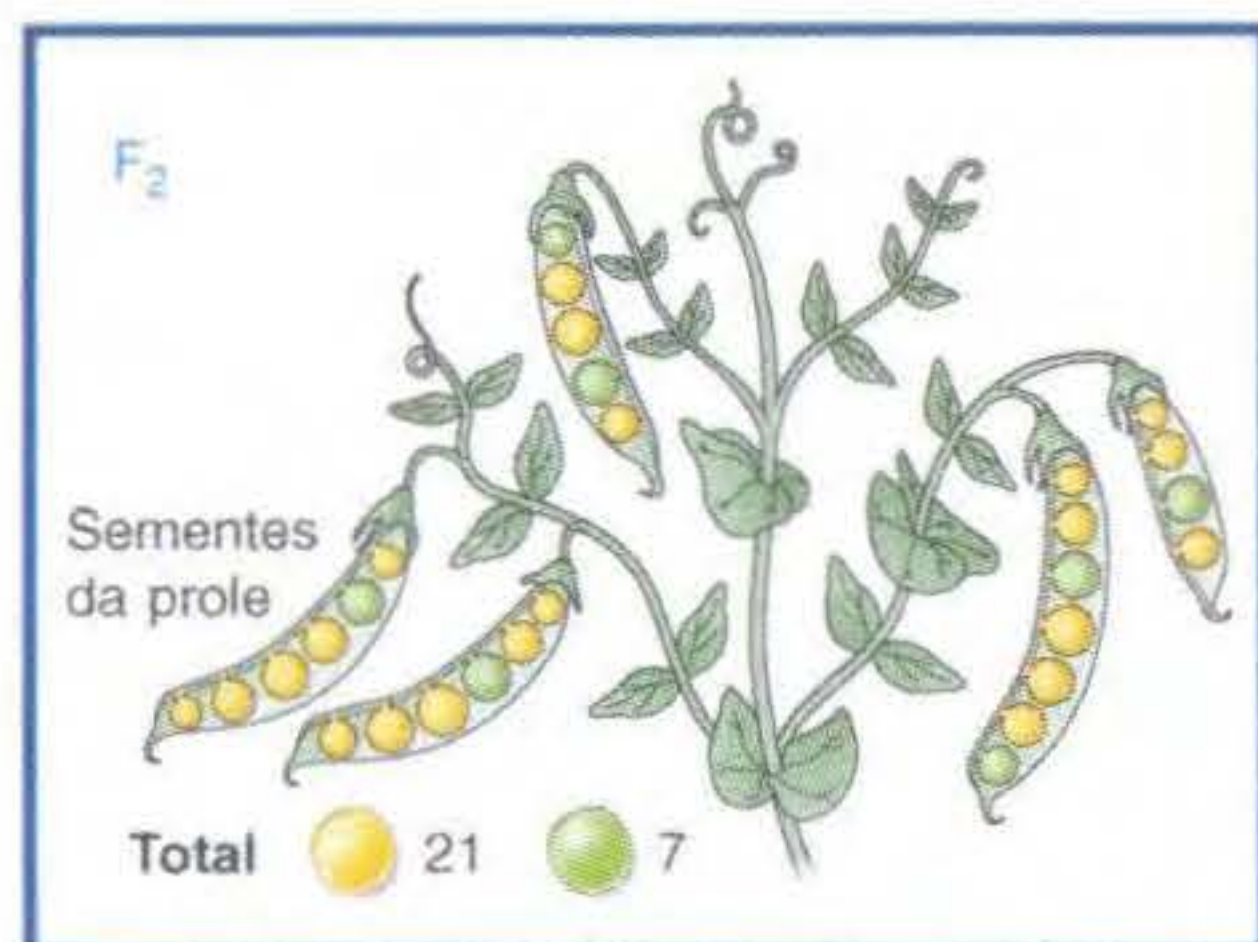
Polimorfismos autossômicos, 58

Distúrbios recessivos ligados ao X, 59

Distúrbios dominantes ligados ao X, 61

Herança ligada ao Y, 62

Cálculo de riscos na análise de heredogramas, 63



3 Distribuição Independente de Genes, 75

3.1 Lei de Mendel da Distribuição Independente, 76

3.2 Trabalhando com a Distribuição Independente, 79

Prevendo as proporções na prole, 80

Uso do teste do qui-quadrado nas proporções monoíbridas e diíbridas, 81

Síntese de linhagens puras, 83

Vigor híbrido, 84

3.3 A Base Cromossômica da Distribuição Independente, 85

Distribuição independente em organismos diplóides, 85

Distribuição independente em organismos haplóides, 86

Distribuição independente de genes autossômicos e ligados ao X, 86

Recombinação, 89

Boxe de Organismo-modelo – *Neurospora*, 89

3.4 Herança Poligênica, 91

3.5 Genes de Organelas: Herança Independente do Núcleo, 93

Padrões de herança em organelas, 95

Segregação citoplasmática, 95

Mutações citoplasmáticas em humanos, 97



4

Mapeamento de Cromossomos Eucarióticos por Recombinação, 109

4.1 Diagnóstico de Ligação, 111

Uso de frequência de recombinantes para reconhecer ligação, 111

Como os crossings produzem recombinantes para genes ligados, 112

Simbolismo de ligação e terminologia, 113

Evidência de que o crossing over é um processo de quebra e reunião, 113

Evidência de que o crossing over ocorre no estágio de quatro cromátides, 113

Múltiplos crossings podem incluir mais de duas cromátides, 114

4.2 Mapeamento por Frequência de Recombinação, 114

Unidades de mapa, 115

Cruzamento-teste de três pontos, 117

Dedução da ordem dos genes por inspeção, 119

Interferência, 119

Uso de proporções como diagnóstico, 120

4.3 Mapeamento com Marcadores Moleculares, 120

Polimorfismos de um só nucleotídeo, 122

Mapeamento usando haplótipos de SNP, 123

Polimorfismos de comprimento de seqüências simples, 125

4.4 Mapeamento de Centrômero com Tétrades Lineares, 127

4.5 Uso do Teste do Qui-quadrado para Testar Análise de Ligação, 129

- 4.6 **Uso de Valores Lod para Avaliar a Ligação em Heredogramas Humanos, 131**
- 4.7 **Cômputo de Crossings Múltiplos Não-vistos, 132**
Função de mapa, 132
A Fórmula de Perkins, 133
- 4.8 **Uso de Mapas Baseados em Recombinação em Conjunto com Mapas Físicos, 134**

5 A Genética de Bactérias e seus Vírus, 151

- 5.1 **Trabalhando com Microrganismos, 153**
- 5.2 **Conjugação Bacteriana, 154**
Descoberta da conjugação, 155
Boxe de Organismo-modelo – *Escherichia coli*, 155
Descoberta do fator fertilidade (F), 156
Linhagens Hfr, 157
Mapeamento de cromossomos bacterianos, 161
Plasmídeos F que levam fragmentos genômicos, 163
Plasmídeos R, 164
- 5.3 **Transformação Bacteriana, 167**
Mapeamento cromossômico usando transformação, 167
- 5.4 **Genética de Bacteriófagos, 167**
Infecção de bactérias por fagos, 168
Mapeamento de cromossomos de fagos usando cruzamentos de fagos, 169
- 5.5 **Transdução, 171**
Descoberta da transdução, 171
Transdução generalizada, 171
Transdução especializada, 173
Mecanismo de transdução especializada, 174
- 5.6 **Mapas Físicos e Mapas de Ligação Comparados, 174**



6 Interação Gênica, 187

- 6.1 **Interações de Alelos de um Único Gene: Variações de Dominância, 189**
Dominância completa e recessividade, 189
Dominância incompleta, 190
Co-dominância, 191

Alelos letais recessivos, 192

Boxe de Organismo-modelo – Camundongo, 194

- 6.2 **Interação de Genes em Vias, 195**
Vias biossintéticas em *Neurospora*, 195
Interação gênica em outros tipos de vias, 197
- 6.3 **Interações Gênicas Interferentes, 198**
Definição de conjuntos de genes usando o teste de complementação, 199
Análise de mutantes duplos por mutações aleatórias, 201
- 6.4 **Penetrância e Expressividade, 209**

PARTE II DO DNA AO FENÓTIPO

7 DNA: Estrutura e Replicação, 225

- 7.1 **DNA: O Material Genético, 226**
Descoberta da transformação, 226
Experimento de Hershey-Chase, 227
- 7.2 **A Estrutura do DNA, 228**
A estrutura do DNA antes de Watson e Crick, 229
A dupla hélice, 231
- 7.3 **Replicação Semiconservativa, 234**
Experimento de Meselson-Stahl, 235
A forquilha de replicação, 235
DNA polimerases, 236
- 7.4 **Visão Geral da Replicação do DNA, 238**
- 7.5 **O Replissomo: Uma Marcante Máquina de Replicação, 239**
Deselcoindizando a dupla hélice, 240
Montagem do replissomo: início da replicação, 241
- 7.6 **Replicação em Organismos Eucarióticos, 241**
O replissomo eucariótico, 242
Origens eucarióticas da replicação, 242
Replicação do DNA e o ciclo celular de levedura, 244
Origens de replicação em eucariontes superiores, 244
- 7.7 **Telômeros e Telomerase: Término da Replicação, 245**
Telômeros, câncer e envelhecimento, 246

8 RNA: Transcrição e Processamento, 253

- 8.1 **RNA, 254**
Os primeiros experimentos sugerem um RNA intermediário, 254
Propriedades do RNA, 255
Classes do RNA, 256
- 8.2 **Transcrição, 257**
Visão Geral: DNA como molde para a transcrição, 257
Estágios da transcrição, 258

8.3 Transcrição em eucariontes, 261

Início da transcrição em eucariontes, 262

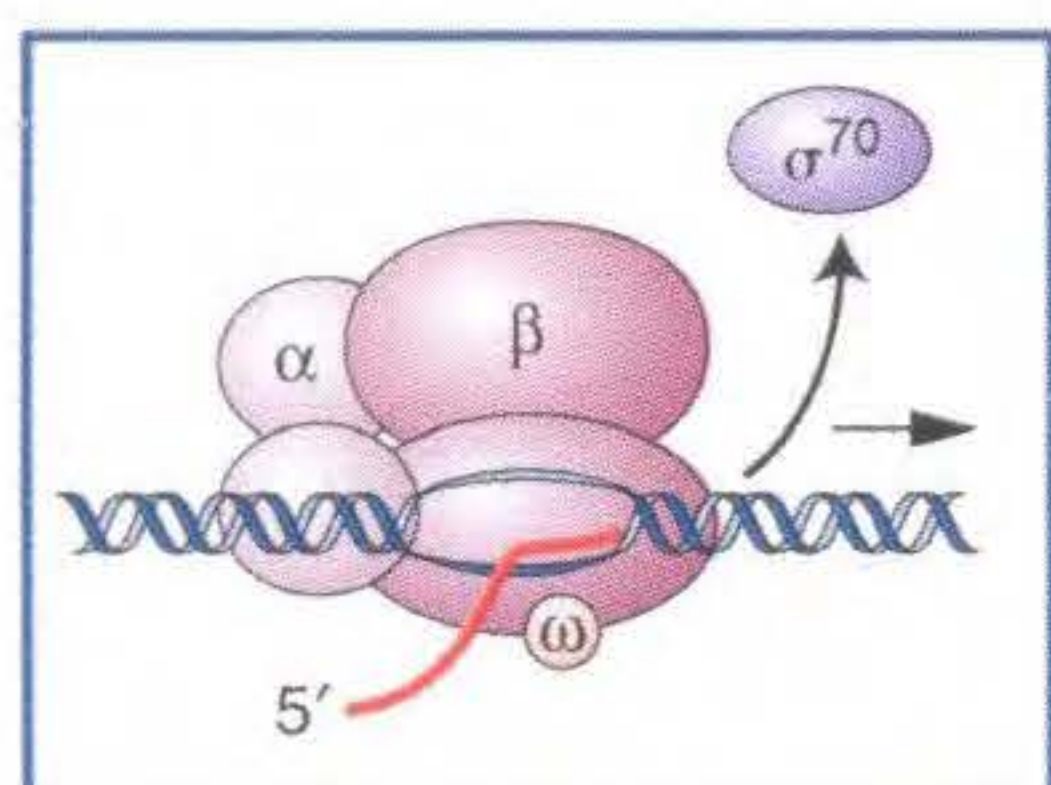
Alongamento, término e processamento pré-mRNA em eucariontes, 263

8.4 RNA Funcionais, 266

Pequenos RNA nucleares (snRNA): o mecanismo de recomposição dos éxons, 266

Auto-recomposição de íntrons e o mundo de RNA, 267

Pequenos RNA de interferência (siRNA), 268

**9 Proteínas e sua Síntese, 275****9.1 Estrutura de Proteínas, 277****9.2 Colinearidade de Gene e Proteína, 279****9.3 O Código Genético, 279**

Códigos superpostos *versus* não-superpostos, 280

Número de letras no códon, 280

Uso de supressores para demonstrar um código triplo, 281

Redundância do código genético, 282

Decifrando o código, 282

Códons de fim, 282

9.4 tRNA: O Adaptador, 283

Tradução do códon pelo tRNA, 283

Novamente a redundância, 285

9.5 Ribossomos, 286

Características dos ribossomos, 286

Início, alongamento e término da tradução, 287

Mutações sem sentido do supressor, 291

9.6 O Proteoma, 291

A recomposição alternativa gera isoformas de proteína, 292

Eventos pós-traducionais, 292

10 Regulação da Expressão Gênica em Bactérias e seus Vírus, 301**10.1 Regulação Gênica, 303**

As bases da regulação transcricional procariótica: interruptores genéticos, 303

Um primeiro exame do circuito regulador *lac*, 304

10.2 Descoberta do Sistema *lac*: Controle Negativo, 306

Genes controlados juntos, 307

Evidência genética do operador e repressor, 307

Evidência genética de alosteria, 309

Análise genética do promotor *lac*, 309

Caracterização molecular do repressor *lac* e operador *lac*, 310

Mutações polares, 311

10.3 Repressão Catabólica do Óperon *lac*: Controle Positivo, 311

A base da repressão catabólica do óperon *lac*: escolher o melhor açúcar para metabolizar, 311

As estruturas dos sítios-alvo de DNA, 312

Um resumo do óperon *lac*, 313

10.4 Controle Duplo Positivo e Negativo: o Óperon de Arabinose, 313**10.5 Vias Metabólicas e Níveis Adicionais de Regulação: Atenuação, 314**

A transcrição do óperon *trp* é regulada em duas etapas, 315

10.6 Ciclos de Vida de Bacteriófagos: Mais Reguladores, Óperons Complexos, 317

Anatomia molecular da mudança genética, 320

Ligação específica de sequência de proteínas regulatórias ao DNA, 321

10.7 Fatores Sigma Alternativos Regulam Grandes Conjuntos de Genes, 322**11 Regulação da Expressão Gênica em Eucariontes, 329****11.1 Regulação da Transcrição em Eucariontes: Visão Geral, 330****Boxe de Organismo-modelo – Leveduras, 333****11.2 Lições de Leveduras: O Sistema GAL, 334**

Gal4 regula vários genes pelas sequências de ativação antecedentes, 334

A proteína Gal4 tem domínios de ligação ao DNA e ativação separáveis, 335

A atividade de Gal4 é regulada fisiologicamente, 336

Funções de Gal4 na maioria dos eucariontes, 336

Ativadores recrutam a maquinaria transcricional, 336

11.3 Cromatina Dinâmica e Regulação Gênica Eucariótica, 337

Proteínas remodeladoras de cromatina e ativação gênica, 338

Histonas e remodelagem de cromatina, 339

11.4 Mecanismo de Ação Acentuadora, 340

O acentuassomo β -interferon, 341

O controle do tipo reprodutivo de levedura: interações combinatórias, 342

As proteínas de ligação ao DNA regulam combinatoriamente a expressão de genes específicos do tipo de célula, 342

Isoladores de bloqueio de acentuador, 342

11.5 Imprinting Genômico, 343

Mas, e quanto a Dolly e outros mamíferos clonados?, 345

11.6 Domínios de Cromatina e sua Herança, 345

Mudança de tipo reprodutivo e silenciamento gênico, 346

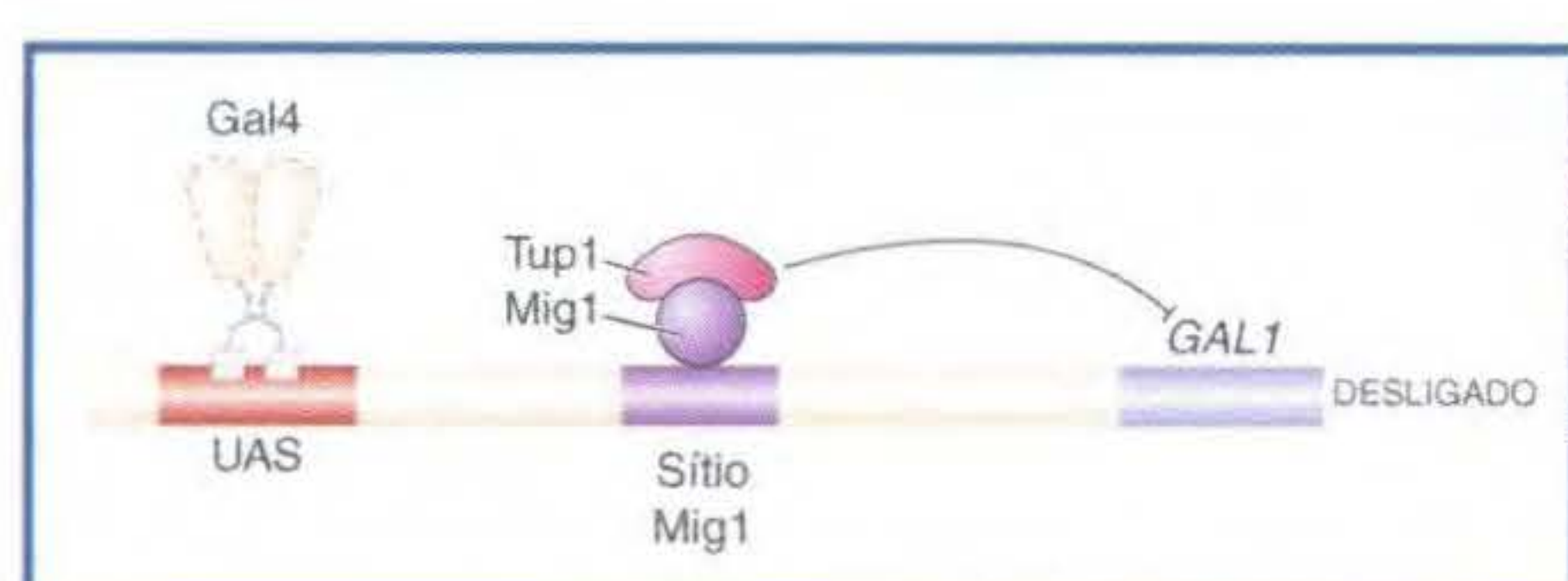
Comparação da heterocromatina e eucromatina, 346

Variação do efeito de posição em *Drosophila* revela a vizinhança genômica, 347

A análise genética de PEV revela proteínas necessárias para a formação de heterocromatina, 348

Silenciamento de um cromossomo inteiro: inativação do cromossomo X, 350

A herança de marcas epigenéticas e estrutura cromossômica, 351



12 O Controle Genético do Desenvolvimento, 355

12.1 O Enfoque Genético do Desenvolvimento, 356

12.2 As Ferramentas para o Desenvolvimento de *Drosophila*, 357

Boxe de Organismo-modelo – *Drosophila*, 358

Classificação dos genes pela função desenvolvimental, 358

Genes homeóticos e identidade segmentar, 359

Organização e expressão dos genes *Hox*, 360

O homeobox, 362

Grupos de genes *Hox* controlam o desenvolvimento na maioria dos animais, 363

12.3 Definição de Todas as Ferramentas, 365

Os eixos ântero-posterior e dorsoventral, 366

Expressão dos genes ferramentais, 366

12.4 Regulação Espacial da Expressão Gênica no Desenvolvimento, 370

Gradientes maternos e ativação gênica, 370

Desenho das faixas: Integração de impulsos de proteínas gap, 372

Tornando os segmentos diferentes: Integração dos estímulos Hox, 373

12.5 Regulação Pós-transcricional da Expressão Gênica no Desenvolvimento, 374

Recomposição do RNA e determinação do sexo em *Drosophila*, 376

Regulação da tradução do mRNA e linhagem celular em *C. elegans*, 377

Controle traducional no embrião inicial, 377

Boxe de Organismo-modelo – *Caenorhabditis elegans*, 378

miRNA controla a época desenvolvimental em *C. elegans* e outras espécies, 379

12.6 Os Vários Papéis dos Genes Ferramentais Individuais, 380

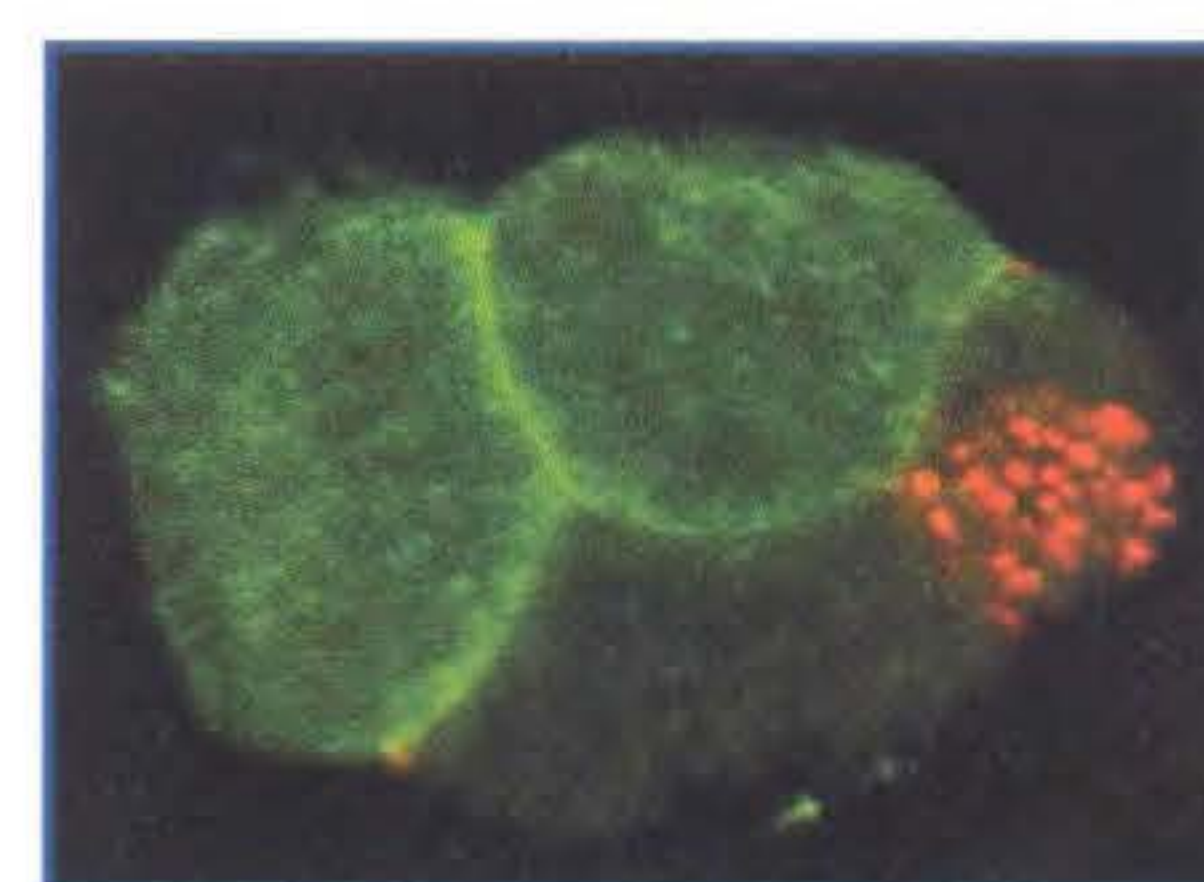
De moscas a dedos, penas e placas do assoalho, 380

12.7 Desenvolvimento e Doença, 381

Polidactilia, 382

Holoprosencefalia, 382

Câncer como uma doença do desenvolvimento, 382



13 Genomas e Genômica, 387

13.1 A Revolução Genômica, 388

13.2 Criando o Mapa de Sequência de um Genoma, 389

Transformando leituras de seqüências em um mapa de seqüência, 389

Estabelecendo uma biblioteca genômica de clones, 391

Seqüenciamento de um genoma simples usando o enfoque shotgun de genoma inteiro, 392

Usando o enfoque shotgun do genoma inteiro para criar uma seqüência-rascunho de um genoma complexo, 392

Usando o enfoque de clone ordenado para seqüenciar um genoma complexo, 394

Preenchendo espaços na seqüência, 395

13.3 Bioinformática: Significado da Seqüência Genômica, 395

A natureza do conteúdo de informação do DNA, 395

Deduzindo os genes codificantes de proteína da seqüência genômica, 395

- 13.4 **A Estrutura do Genoma Humano, 399**
- 13.5 **Genômica Comparativa, 400**
 Camundongos e Humanos, 400
 Genômica comparativa de chimpanzés e humanos, 402
 Elementos conservados e ultraconservados não-codificantes, 402
 Genômica comparativa de *E. coli* não-patogênica e patogênica, 403
- 13.6 **Genômica Funcional e Genética Reversa, 405**
 Ome, *Sweet Ome*, 405
 Genética reversa, 408

PARTE III MUTAÇÃO, VARIAÇÃO E EVOLUÇÃO

14 **O Genoma Dinâmico: Elementos de Transposição, 417**

- 14.1 **Descoberta dos Elementos de Transposição em Milho, 418**
 Os experimentos de McClintock: o elemento *Ds*, 418
Boxe de Organismo-modelo – Milho, 419
 Elementos autônomos e não-autônomos, 421
 Elementos transponíveis: só em milho?, 421
- 14.2 **Elementos de Transposição em Procariontes, 422**
 Sequências bacterianas de inserção, 422
 Transposons procarióticos, 423
 Mecanismo de transposição, 423
- 14.3 **Elementos de Transposição em Eucariontes, 425**
 Classe I: retrotransposons, 425
 Transposons de DNA, 428
 Utilidade dos transposons de DNA para a descoberta de gene, 431
- 14.4 **O Genoma Dinâmico: Mais Elementos de Transposição do que se Imaginava, 431**
 Genomas grandes são amplamente elementos de transposição, 432
 Elementos de transposição no genoma humano, 433
 As gramíneas: retrotransposons LTR desenvolvem-se em grandes genomas, 434
 Abrigos seguros, 434

15 **Mutação, Reparo e Recombinação, 439**

- 15.1 **As Consequências Fenotípicas das Mutações no DNA, 440**
 Tipos de mutação de ponto, 440
 As consequências moleculares das mutações de ponto em uma região codificante, 442
 As consequências moleculares das mutações de ponto em uma região não-codificante, 442

15.2 **A Base Molecular das Mutações Espontâneas, 443**

Teste de flutuação de Luria e Delbrück, 443
 Mecanismos de mutações espontâneas, 445
 Mutações espontâneas em humanos: doenças de repetições de trinucleotídeos, 448

15.3 **A Base Molecular das Mutações Induzidas, 449**

Mecanismos de mutagenese, 449
 O teste de Ames: avaliação de mutágenos em nosso ambiente, 451

15.4 **Mecanismos Biológicos de Reparo, 453**

Reversão direta de DNA danificado, 454
 Reparo por excisão de base, 454
 Reparo por excisão de nucleotídeo, 456
 Reparo pós-replicação: reparo de mal pareamento, 458
 Reparo propenso a erro: síntese de DNA translesão, 459

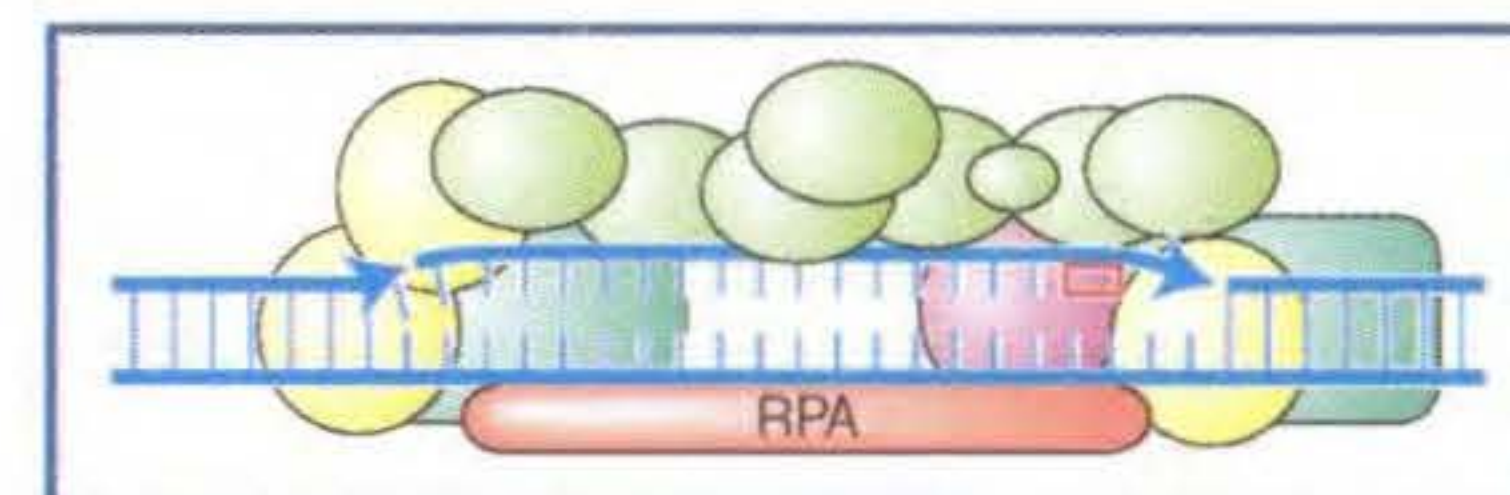
Reparo de quebras bifilamentares, 461

15.5 **O Mecanismo de Crossing Over Meiótico, 463**

Quebras bifilamentares programadas iniciam a recombinação meiótica, 463
 A análise genética de tétrades fornece indícios do mecanismo de recombinação, 464
 O modelo de quebra bifilamentar para recombinação meiótica, 465

15.6 **Câncer: Uma Consequência Fenotípica Importante da Mutação, 467**

Como as células cancerosas diferem das células normais, 467
 Mutações em células cancerosas, 467



16 **Alterações Cromossômicas em Larga Escala, 475**

- 16.1 **Mudanças no Número de Cromossomos, 476**
 Euploidia aberrante, 476
 Aneuploidia, 484
 O conceito de balanço gênico, 488
- 16.2 **Mudanças na Estrutura de Cromossomos, 490**
 Deleções, 492
 Duplicações, 494
 Inversões, 495
 Translocações recíprocas, 499

Translocações robertsonianas, 499
 Aplicações de inversões e translocações, 500
 Rearranjos e câncer, 502
 Identificação das mutações cromossômicas pela genômica, 503

16.3 Incidência Geral de Mutações Cromossômicas Humanas, 503



17 Genética de Populações, 517

17.1 Variação e sua Modulação, 518

Observações de variação, 518
 Polimorfismos de proteína, 520
 Estrutura do DNA e polimorfismo de seqüências, 521

17.2 Efeitos da Reprodução Sexual na Variação, 525

Segregação meiótica e equilíbrio genético, 525
 Heterozigose, 527
 Reprodução aleatória, 527
 Endogamia e reprodução preferencial, 528

17.3 Fontes de Variação, 530

Variação por mutação, 530
 Variação por recombinação, 531
 Variação por migração, 532

17.4 Seleção, 533

Duas formas de seleção, 533
 Medindo as diferenças de adaptabilidade, 534
 Como a seleção funciona, 534
 Taxa de mudança em frequência gênica, 535

17.5 Polimorfismo balanceado, 537

Sobredominância e subdominância, 537
 Balanço entre mutação e seleção, 538

17.6 Eventos Aleatórios, 538

18 Genética Quantitativa, 545

18.1 Genes e Características Quantitativas, 546

18.2 Algumas Noções Estatísticas Básicas, 548

Distribuições estatísticas, 548
 Medidas estatísticas, 548

18.3 Genótipos e Distribuição Fenotípica, 549

A diferença crítica entre características quantitativas e mendelianas, 549
 Número de genes e características quantitativas, 550

18.4 Norma de Reação e Distribuição Fenotípica, 551

18.5 Determinando as Normas de Reação, 552

Plantas e animais domésticos, 553
 Estudos de populações naturais, 554
 Resultados dos estudos de normas de reação, 554

18.6 A Herdabilidade de uma Característica Quantitativa, 556

Familialidade e herdabilidade, 556
 Similaridade fenotípica entre parentes, 557

18.7 Quantificando a Herdabilidade, 557

Métodos para estimar H^2 , 558
 O significado de H^2 , 559
 Herdabilidade restrita, 561
 Estimando os componentes da variância genética, 562
 Seleção artificial, 563
 O uso de h^2 em endogamia, 564

18.8 Localizando os Genes, 565

Segregação de genes marcadores, 566
 Análise de ligação quantitativa, 567

Apêndice Estatístico, 568

Medidas da tendência central, 568
 Medidas de dispersão: a variância, 570
 Medidas de relação, 570

19 Genética Evolutiva, 579

19.1 Evolução Darwiniana, 580

19.2 Uma Síntese de Forças: Variação e Divergência de Populações, 581

19.3 Vários Picos Adaptativos, 584

Exploração dos picos adaptativos, 585

19.4 Variação Genética, 586

Herdabilidade da variação, 586
 Variação dentro de populações e entre populações, 588

19.5 Mutação e Evolução Molecular, 588

Sinal característico da seleção purificadora no DNA, 588

19.6 Relação entre Genética e Mudança Funcional na Evolução de Proteínas, 590

Sinal característico da seleção positiva nas seqüências de DNA, 590

Evolução morfológica, 590

Inativação gênica, 592

19.7 Evolução Regulatória, 593

Evolução regulatória em humanos, 595

19.8 A Origem de Novos Genes, 595

Poliploidia, 595

Duplicações, 595

DNA importado, 597

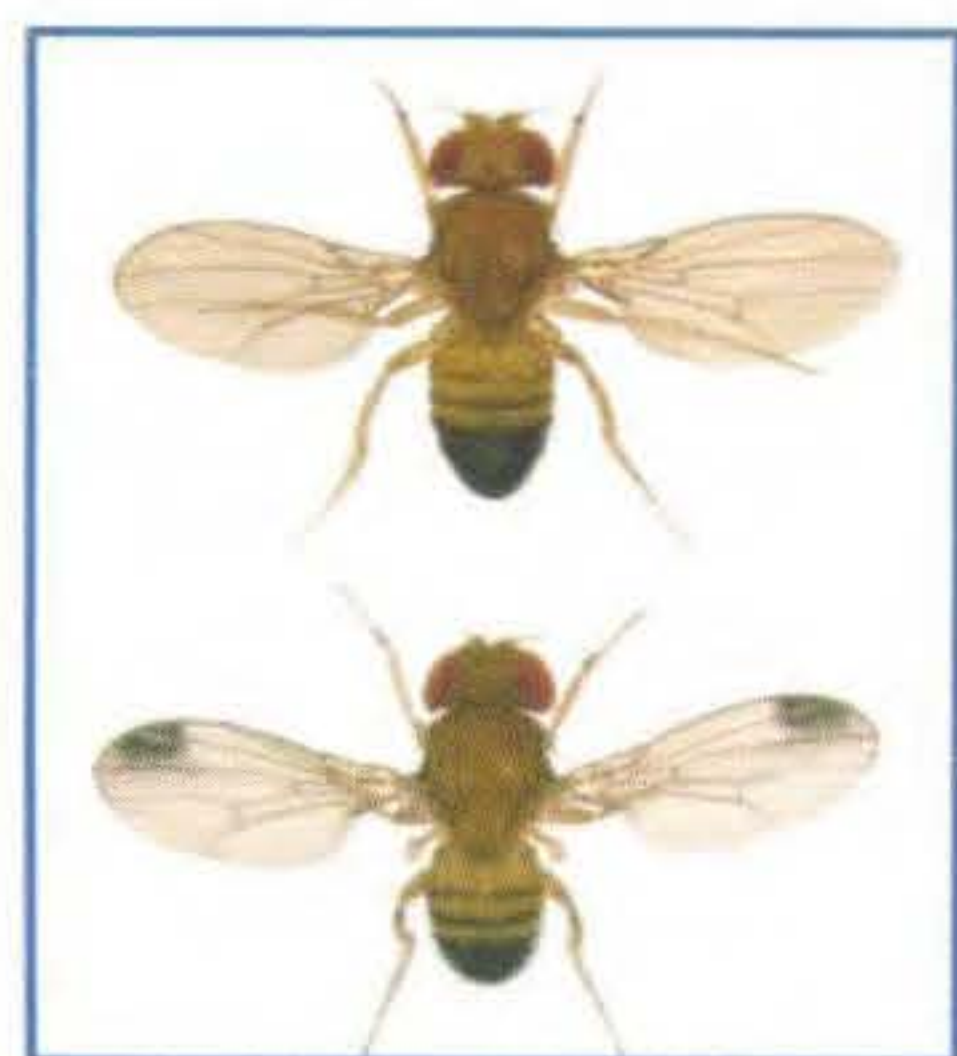
19.9 Evidência Genética de Ancestralidade Comum na Evolução, 598

Comparando os proteomas entre espécies distantes, 600

Comparação dos genomas entre vizinhos próximos: Genômica comparativa humano-camundongo, 601

19.10 O Processo de Especiação, 601

Genética do isolamento de espécies, 603



PARTE IV TÉCNICAS

20 Isolamento e Manipulação de Gene, 609

20.1 Gerando Moléculas de DNA Recombinante, 610

Tipo de DNA doador, 611

Cortando o DNA genômico, 611

Ligação do DNA doador e vetor, 612

Amplificação dentro de uma bactéria, 612

Entrada de moléculas recombinantes nas bactérias, 615

Recuperação de moléculas recombinantes amplificadas, 615

Construção de bibliotecas genômicas e de cDNA, 616

Encontrar um clone específico de interesse, 617

20.2 Amplificação de DNA *in Vitro*: a Reação em Cadeia da Polimerase, 622

20.3 Determinando a Sequência de Bases de um Segmento de DNA, 624

20.4 Análise Genética Direta Usando Clonagem Posicional, 626

Uma análise direta para identificar um gene de doença humana, 627

Uma análise direta para identificar um gene importante para a domesticação do milho, 628

20.5 Detecção de Alelos de Doenças Humanas: Diagnóstico Genético Molecular, 629

20.6 Engenharia Genética, 631

Engenharia genética em *Saccharomyces cerevisiae*, 631

Engenharia genética em plantas, 631

Engenharia genética em animais, 633

Terapia gênica humana, 639

Um Guia Resumido para Organismos-modelo, 647

Apêndice A: Nomenclatura Genética, 663

Apêndice B: Recursos de Bioinformática para Genética e Genômica, 665

Glossário, 667

Respostas de Problemas Seleccionados, 689

Índice Alfabético, 701

Prefácio

Nas últimas décadas, o poder da análise genética cresceu com a introdução da tecnologia do DNA recombinante e, então, a genômica. Nesta nova edição do *Introdução à Genética*, nossa meta é comunicar aos estudantes o poder da análise genética para iluminar todos os aspectos da biologia. Fizemos várias análises substanciais no conteúdo e na organização do livro para melhor transmitir a prática atual da genética.

Sean Carroll Junta-se ao Grupo de Autores

Um novo co-autor, Sean B. Carroll, juntou-se ao grupo de autores. O Dr. Carroll é Pesquisador no Howard Hughes Medical Institute e Professor de Biologia Molecular e Genética na University of Wisconsin, em Madison, onde também leciona genética e biologia desenvolvimental. É pesquisador e professor no campo da biologia desenvolvimental. Seus *insights* em desenvolvimento evolutivo, genômica comparativa e regulação gênica tornam claro como o estudo da genética está transformando nossa compreensão da evolução, do desenvolvimento e da biologia molecular.

Mais Ênfase na Prática da Genética de Hoje

Nova Ênfase na Aplicação da Análise Genética a uma Ampla Gama de Processos Biológicos

Nossa meta é mostrar, com vários exemplos novos, como a identificação de genes e suas interações é uma ferramenta importante para responder a questões biológicas de muitos tipos, desde “O que torna algumas linhagens de *E. coli* patogênicas?” a “Como os americanos nativos domesticaram o milho?” Tanto nos primeiros capítulos sobre transmissão quanto nos capítulos moleculares, o texto demonstra como a análise genética fornece um enfoque sistemático ao estudo de toda a biologia. Nos capítulos de transmissão, o material sobre genética mendeliana foi reestruturado para colocar os resultados de Mendel diretamente no contexto da pesquisa da genética moderna e o processo da descoberta do gene. Desde o começo, o estudante segue o processo da dissecação genética, começando com a identificação monogênica no Capítulo 2, por meio do mapeamento genético no Capítulo 4 até as interações gênicas no Capítulo 6. Os capítulos moleculares cerne ampliam a estratégia desenvolvida nos capítulos de transmissão para mostrar como um gene é isolado e sua função é analisada, destacando as descobertas modernas e as práticas.

Nova Ênfase nas Aplicações Humanas Recentes

Novos exemplos humanos de pesquisas recentes foram adicionados para despertar o interesse do estudante e demonstrar a relevância da genética:

- O uso do *hapmap* humano para encontrar esses genes (Capítulo 4)
- Como as interações gênicas determinam a cor do cabelo humano (Capítulo 6)
- Osteogênese imperfeita (doença dos ossos quebradiços) como um exemplo de um dominante negativo (Capítulo 6)
- Telômeros, câncer e envelhecimento (Capítulo 7)
- Ação de antibióticos e ribossomos (Capítulo 9)
- Nova seção importante sobre desenvolvimento e doença: polidactilia, holoprosencefalia, câncer como uma doença do desenvolvimento (Capítulo 12)
- Seção ampliada sobre a estrutura do genoma humano (Capítulo 13)

- Novas seções de genômica comparativa em camundongos e humanos e a genômica comparativa de chimpanzés e humanos (Capítulo 13).
- Como foi identificado um elemento codificante para o gene humano *ISL1* (Capítulo 13)
- Xeroderma pigmentoso e síndrome de Cockayne, ambos causados por defeitos no reparo por excisão de nucleotídeos (Capítulo 15)
- Evolução regulatória em humanos: antígeno Duffy em hemácias (Capítulo 19)
- Determinação dos genes que foram selecionados na domesticação do milho há 10.000 anos (Capítulo 20)

Ampla Cobertura de Genômica Comparativa

No capítulo revisado de genômica, vimos em profundidade como a genômica comparativa informa a análise genética e revela diferenças cruciais entre os organismos, respondendo a perguntas como “O que torna uma linhagem virulenta de *E. coli* tão perigosa?” e “Qual a diferença genética entre humanos e chimpanzés?” Nós então examinaremos os métodos emergentes para trabalhar a partir de uma sequência genômica para a análise do funcionamento gênico. O capítulo destaca como a genômica está mudando o modo como os geneticistas trabalham.

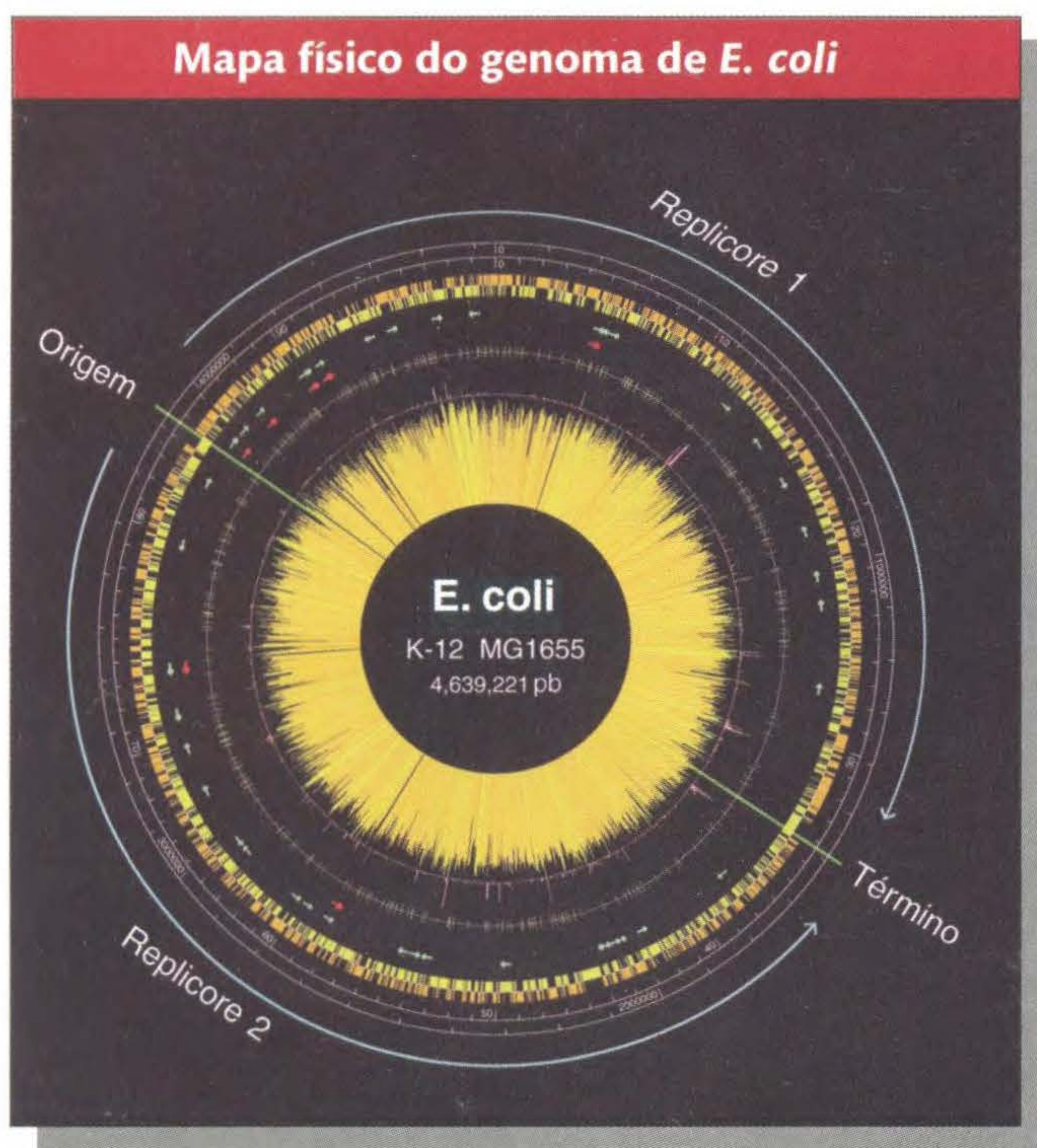


FIG. 5.36 Esse mapa foi obtido do sequenciamento do DNA e posições de plotagem de genes. [*Science* 277, 1997, 1453-1462. Imagem por cortesia do Dr. Guy Plunkett III. Reproduzido de AAAS, com permissão.]

Novas Abordagens Demonstrando como a Genética é Praticada Hoje em Dia

Uma nova abordagem chamada “O que os Geneticistas Estão Fazendo Hoje” relata como as técnicas genéticas estão sendo usadas hoje para responder a dúvidas biológicas específicas, tais como “Qual a ligação entre o encurtamento do telômero e o envelhecimento?” ou “Como podemos encontrar componentes ausentes em uma via biológica específica?”

Experimentos de Ponta

Sobre o foco tradicional do texto nos experimentos clássicos, os capítulos moleculares agora apresentam evidências e considerações que levaram a alguns avanços importantes, tais como a descoberta de um novo modelo de recombinação e a acentuada compreensão da regulação gênica eucariótica, entre outros.

Novo Tratamento do Desenvolvimento

Os conceitos de desenvolvimento fundamentam-se em uma história contínua, começando com capítulos de regulação gênica e terminando com um capítulo mais focado no desenvolvimento. Esse capítulo usa o desenvolvimento como um modelo para demonstrar como o estudo das mutações revela os componentes de uma via biológica.

Nova Seção Ilustrando o Processo Completo de uma Análise Genética Avançada

Uma nova seção no Capítulo 20 apresenta dois estudos de casos mostrando como as técnicas e estratégias da análise genética são usadas para isolar um gene e, portanto, resolver um problema biológico. Os dois casos incluem um novo estudo de caso isolando o gene que tornou possível a domesticação do milho.

Cobertura Mais Ampla de Tópicos Importantes

Cobertura Mais Ampla de Regulação Gênica

Nesta nona edição, ampliamos a cobertura da regulação gênica com mais dois capítulos: um sobre mecanismos bacterianos e outro sobre mecanismos eucarióticos. Essa mudança corresponde ao alargamento de nossa compreensão de importantes mecanismos genéticos reguladores que são únicos aos eucariontes. O novo capítulo sobre regulação gênica em eucariontes destaca o papel da estrutura da cromatina em controlar a expressão gênica e apresentar o trabalho experimental mais importante que iluminou os mecanismos de controle particulares aos eucariontes.

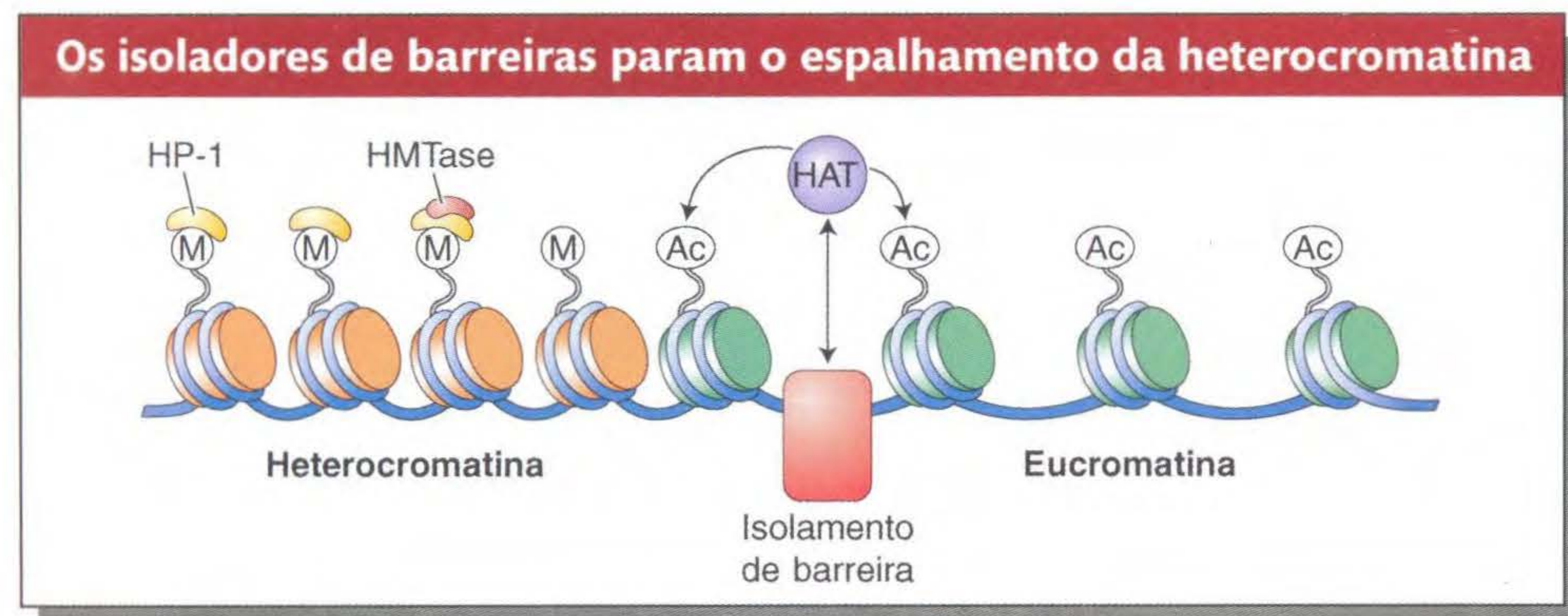


FIG. 11.26 Nesse modelo, os isoladores de barreiras recrutam atividades enzimáticas, tais como histona acetiltransferase (HAT), que promovem a formação de eucromatina. A letra "M" representa metilação, e as letras "Ac", acetilação. [De M. Gaszner and G. Felsenfeld. "Insulators: Exploiting Transcriptional and Epigenetic Mechanisms," *Nat. Rev. Genet.* 7, 2006, 703-713.]

Cobertura Mais Ampla do Reparo do DNA e Recombinação

Novos *insights* da base molecular das doenças humanas estimularam novas pesquisas sobre os mecanismos subjacentes ao reparo do DNA em animais. Essa nova pesquisa nos fez ampliar e atualizar a cobertura do reparo do DNA e recombinação na nona edição, trazendo essas novas e relevantes descobertas aos estudantes de graduação.

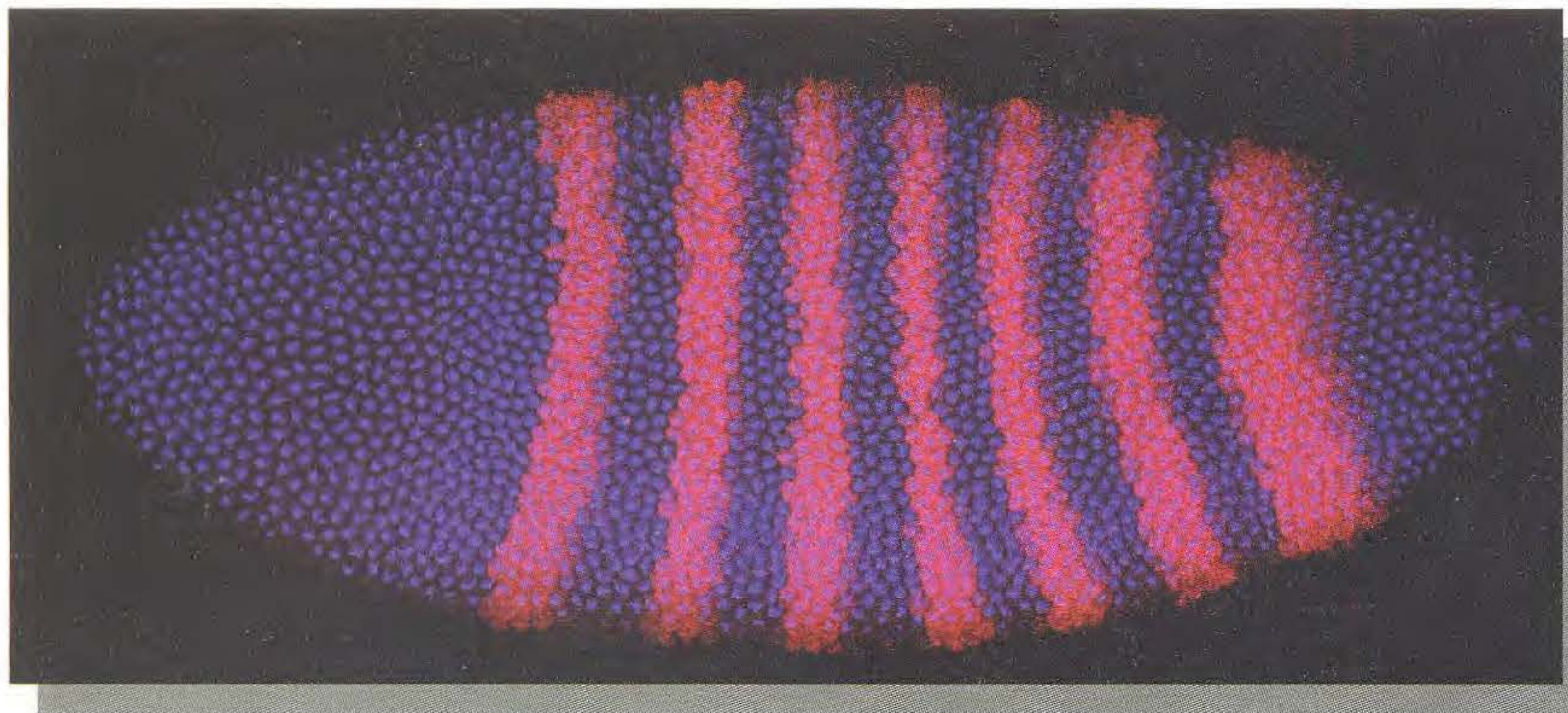
Nova Organização

A Genética Mendeliana é Coberta em Dois Capítulos

A genética mendeliana, tradicionalmente tratada em um capítulo, foi dividida em dois capítulos essencialmente baseados na primeira e segunda leis de Mendel, descrevendo a herança monogênica e a distribuição independente. Essa divisão permite aos estudantes ver claramente os dois processos como os dois pilares principais da genética com aplicações um pouco diferentes. Como parte desse realinhamento, o material da teoria cromossômica da hereditariedade foi distribuído pelos capítulos, e aquele específico sobre o assunto foi suprimido. Entretanto, a essência da análise histórica que estabeleceu a teoria cromossômica ainda está presente.

O Desenvolvimento Segue o Tratamento da Regulação Gênica

Essa ordem é mais lógica porque os conceitos de regulação gênica, aprendidos nos Capítulos 10 e 11 dos modelos primariamente unicelulares, formam uma base sólida para aprender como a regulação gênica é controlada no espaço e ao longo do tempo nos animais multicelulares. A nova colocação fornece a maior continuidade entre as idéias centrais em cada capítulo. Por exemplo, ela conecta os modelos mais simples de mudanças genéticas e proteínas regulatórias nas bactérias às mudanças mais complexas que controlam a formação de planos corpóreos e partes do corpo em animais.



Pág. 355 Expressão gênica de um desenvolvimento de embrião da mosca-das-frutas. As sete faixas vermelhas marcam as células que expressam o mRNA de um gene codificante de uma proteína reguladora que controla o número de segmentos no embrião de *Drosophila*. A regulação espacial da expressão gênica é central para o controle do desenvolvimento animal. [Foto de Dave Kosman, Ethan Bier and Bill McGinnis.]

A Cobertura do Câncer Agora é Integrada

Como parte da ênfase em como a genética funciona, cobrimos o câncer através de um exemplo dos vários mecanismos genéticos subjacentes a ele. Assim, esses mecanismos do câncer são colocados em seus contextos apropriados, e não há capítulo que lide especificamente com os processos heterogêneos subjacentes ao câncer.

As Técnicas da Tecnologia do DNA Agora Estão no Capítulo Final

Nessa posição, as técnicas não comprometem o fluxo dos capítulos do núcleo molecular, embora a informação esteja em um lugar onde pode ser facilmente acessada.

Avanços Recentes e Outros Acréscimos

A nona edição foi totalmente atualizada, incluindo novas discussões sobre os seguintes avanços recentes e outros tópicos:

- Mapeamento usando haplótipos de SNP (Capítulo 4)
- Uso de mapas baseados em recombinação em conjunto com mapas físicos (Capítulo 4)
- RNA de interferência: a história de sua descoberta, seu mecanismo e como ele causa o silenciamento transgênico (Capítulo 8)
- Nova seção sobre proteoma, incluindo novo material sobre
 - Recomposição alternativa e isoformas de proteínas
 - Fosforilação e ubiquitinação de proteínas (Capítulo 9)
- Atenuação (Capítulo 10)
- Regulação do ciclo de vida do fago λ (Capítulo 10)

- Regulação por fatores sigma alternativos (Capítulo 10)
- O sistema GAL, incluindo os domínios separáveis, regulação fisiológica, recrutamento da maquinaria transcricional por ativadores, e coativadores (Mediador) (Capítulo 11)
- O controle do tipo reprodutivo em leveduras: interações combinatórias (Capítulo 11)
- Isoladores de bloqueio de acentuador (Capítulo 11)
- Mais sobre como funciona o imprinting em âmbito molecular, incluindo o papel da metilação do DNA e como o espalhamento da heterocromatina é parado pela barreira de insuladores (Capítulo 11)
- Recomposição do RNA e determinação do sexo em *Drosophila* (Capítulo 12)
- Nova seção lidando com o controle traducional no embrião inicial, a união de proteínas de ligação ao mRNA para reprimir a tradução e estabelecer destinos celulares distintos (Capítulo 12)
- Controle pelo miRNA da época desenvolvimental (Capítulo 12)
- Genômica comparativa: genoma de camundongo *versus* humano; comparação de chimpanzés e humanos; identificação de regiões não-codificantes funcionais procurando seqüências conservadas; bactérias patogênicas *versus* não-patogênicas (Capítulo 13)
- Dois tipos de reparo por excisão de nucleotídeos: reparo genômico global e reparo por excisão de nucleotídeo acoplado a transcrição (Capítulo 15)
- Discussões atualizadas de várias vias de reparo e ligações a doenças humanas (Capítulo 15)
- Discussão sobre recombinação meiótica totalmente reescrita, atualizada e expandida (Capítulo 15)
- Nova seção importante sobre a evolução de proteínas (Capítulo 19)
- Nova seção importante de evolução por meio de mudanças nos elementos regulatórios (Capítulo 19)

Características Continuadas

Cobertura de Organismos-modelo

A nona edição conserva a acentuada cobertura dos sistemas-modelo em formatos práticos e flexíveis tanto para estudantes quanto para professores.

- O Capítulo 1 inclui uma seção sobre organismos-modelo.
- Boxes de Organismo-modelo, apresentados no contexto apropriado, dão informações adicionais sobre o organismo na natureza e seu uso experimentalmente.
- Um Guia Resumido de Organismos-modelo, ao final do livro, fornece um acesso rápido a informações essenciais, práticas, sobre o uso de organismos-modelo específicos em pesquisas.
- Um Índice de Organismos-modelo, no lado interno da capa, fornece referências a páginas, capítulo por capítulo, para discussões sobre organismos específicos no texto, permitindo aos professores e estudantes achar facilmente e reunir informação comparável nos organismos.

Grupos de Problemas

Independentemente da clareza da exposição, a compreensão profunda requer que o estudante se envolva pessoalmente com o material. Daí nossos esforços



FIG. 8.19 Capa da revista *Science* de dezembro de 2002. [De *Science*, vol. 298, no. 5602, December 20, 2002. Reproduzido com permissão da AAAS.]

para estimular o estudante a resolver os problemas. Com foco na análise genética, a nona edição oferece aos estudantes oportunidades para praticar as habilidades em resolver problemas, tanto no texto quanto online, por meio das seguintes características:

- **Conjuntos Versáteis de Problemas.** Os problemas cobrem toda uma gama de graus de dificuldade. Eles são categorizados de acordo com o nível de dificuldade, básico ou desafiador.
- **Problemas Resolvidos.** Ao final de cada capítulo, esses exemplos desenvolvidos ilustram como os geneticistas aplicam princípios a dados experimentais.
- **Solucionando o Problema.** Um problema de genética contém uma matriz complexa de conceitos e informações. “Solucionando o Problema” ajuda os estudantes a aprender como abordar estrategicamente a solução do problema, uma etapa de cada vez, conceito a conceito.
- **Explorando Genomas: Tutoriais de Bioinformática Baseados na Web (Exploring Genomes: Web-Based Bioinformatics Tutorials), 2nd ed.** Esses tutoriais, disponíveis apenas na edição em inglês pela Editora W.H. Freeman, no site Student Web em <http://www.whfreeman.com/iga9e>, guia os estudantes por procuras e análises no banco de dados do National Center for Biotechnology Information (NCBI). As atividades introduzem os estudantes às bases do BLAST, Entrez, Pub-Med, OMIM, e mais. Esta edição inclui três novos tutoriais: o banco de dados COG, análise funcional e genômica ambiental.

Mídia e Suplementos (disponíveis apenas na edição em inglês pela Editora W.H. Freeman)

PARA OS PROFESSORES

MÍDIA (apenas na edição em inglês)

NOVO eBook ISBN: 1-4292-0453-2 O eBook integra totalmente o texto e é uma mídia interativa em um formato que apresenta uma variedade de ferramentas úteis de estudo (texto completo, pesquisa estilo Google, notas, marcação do livro, destaques, e mais).

NOVO Clicker Questions (Questões Clicker) (apenas na edição em inglês)

Discussões de início rápido iluminam pontos importantes, promovendo uma melhor compreensão conceitual durante as leituras. Escrito por Sylvia Fromherz, University of Colorado at Boulder. Para professores iniciantes na tecnologia clicker, a faculdade pode se inscrever para seções de treinamento WebEx em www.whfreeman.com/iga9e.

NOVO Layered PowerPoints (Etapas em PowerPoint) (apenas na edição em inglês)

Iluminam tópicos difíceis para os estudantes, analisando conceitos genéticos complexos e processos etapa por etapa em um formato visual. Preparado por Wayne Forrester, Indiana University.

Todas as Imagens do Texto (apenas na edição em inglês)

Acesso a mais de 500 ilustrações em vários formatos para download. PowerPoints e JPEG de alta resolução com legendas aumentadas para projetar as imagens mais claramente para apresentações em aulas. Adicionalmente, esses arquivos de PowerPoint e JPEG estão disponíveis sem legendas para fácil padronização em PowerPoint.

45 Animações em FLASH em Etapas e Contínuas (totalmente em inglês)

Essas animações foram criadas por Anthony Griffiths em conjunto com BioStudio Visual Communications. A lista completa de animações aparece no final deste Prefácio.

Banco de Acesso (apenas na edição em inglês)

Esse recurso junta uma ampla seleção de problemas de genética para uso em testes, tarefas para casa ou atividades em classe. Localizável por tópico e fornecido no formato MS Word, o banco oferece um alto nível de flexibilidade. Escrito por Ikhide Imumorin, Spelman College.

Manual de Soluções para o Estudante (apenas na edição em inglês)

Escrito por William Fixsen, Harvard University, e Diane K. Lavett, Georgia Institute of Technology, o *Solutions Manual* contém soluções desenvolvidas para todos os problemas no texto, incluindo os exercícios “Solucionando o Problema”. Todas as soluções foram recentemente verificadas, quanto à exatidão, por James Price, do Utah Valley State College. Disponível no site Instructor’s Web como arquivos de Word fáceis de imprimir.

Instructor’s Resource CD-ROM (apenas na edição em inglês)

ISBN: 1-4292-0176-2

Contém todas as imagens do texto em PowerPoint e JPEG, 45 FLASH Animations, Clicker Questions, Layered PowerPoints e Assessment Bank.

Password-Protected Instructor’s Web Site at

www.whfreeman.com/iga9e (apenas na edição em inglês)

Inclui todos os recursos eletrônicos citados aqui.

SUPLEMENTOS

Conjunto de Transparências (apenas na edição em inglês)

ISBN: 1-4292-0175-4

O conjunto em cores de transparências contém 150 ilustrações-chave do texto com legendas aumentadas que são claramente projetadas em uma palestra em sala de aula.

Understanding Genetics: Strategies for Teachers and Learners in Universities and High Schools (apenas na edição em inglês)

ISBN: 0-7167-5216-6

Escrito por Anthony Griffiths e Jolie-Mayer Smith, essa coleção de artigos enfoca a solução de problemas e descreve os métodos para ajudar os estudantes a melhorar suas habilidades em processar e integrar novas informações.

PARA OS ESTUDANTES

MÍDIA

Student Web Site em www.whfreeman.com/iga9e (totalmente em inglês)



Quarenta e cinco animações FLASH estão disponíveis em vídeo para melhor transporte e flexibilidade. Os estudantes podem agora rever processos e conceitos genéticos importantes, de acordo com sua conveniência, fazendo download das animações em seus aparelhos de MP3.

- **Online Sample Tests (apenas na edição em inglês).** Os estudantes podem testar sua compreensão e receber um retorno imediato respondendo às perguntas online que cobrem os conceitos-núcleo em cada capítulo. As questões têm a referência da página do texto para uma fácil revisão do material.
- **Step-Through and Continuous Play FLASH Animations (totalmente em inglês).** Essas animações foram criadas por Anthony Griffiths em conjunto com o BioStudio Visual Communications. A lista completa das animações aparece na página de face.
- **Interactive “Unpacking the Problem” (apenas na edição em inglês)** Um exercício para o conjunto de problemas da maioria dos capítulos está disponível online de forma interativa. Como na versão do texto, cada “Solucionando o Problema” da web usa uma série de questões para que os estudantes gradativamente entendam os processos para resolver o problema. A versão online oferece um *feedback* imediato aos estudantes à medida que trabalham os problemas, bem como um caminho e gradação convenientes. Escrito por Craig Berezowsky, University of British Columbia.
- **NEW Exploring Genomes: Web-Based Bioinformatics Tutorials, 2nd ed. (disponíveis apenas na edição em inglês pela Editora W.H. Freeman)** Paul G. Young, da Queen’s University, Kingston, Ontario, atualizou seus tutoriais interativos, que ensinam aos estudantes as bases para usar corretamente os vastos recursos genômicos disponíveis no National Center for Biotechnology Information (NCBI) Web site. Os tutoriais guiam os estudantes pelos exercícios, incluindo procuras ao vivo para ilustrar o potencial desses recursos de bioinformática para aplicação na análise genética moderna. Além disso, esta edição inclui três novos tutoriais cobrindo genômica ambiental, o banco de dados COG e análise funcional.

SUPLEMENTOS

Solutions MegaManual (apenas na edição em inglês)

ISBN: 1-4292-0177-0

- **Solutions Manual.** Escrito por William Fixsen, Harvard University, e Diane K. Lavett, Georgia Institute of Technology, o *Solutions Manual* contém soluções desenvolvidas para todos os problemas no texto, incluindo os exercícios de “Solucionando o Problema”. Usado em conjunto com o texto, esse manual é um dos melhores meios para desenvolver uma ampla compreensão dos princípios de genética. Todas as soluções foram recentemente verificadas, quanto à precisão, por James Price, do Utah Valley State College.
- **NOVO Exploring Genomes: Web-Based Bioinformatics Tutorials, 2nd ed.** Escrito por Paul G. Young, da Queen’s University, Kingston, Ontario, esses tutoriais interativos são usados em conjunto com os tutoriais online encontrados em www.whfreeman.com/iga9e para orientar os estudantes em pesquisas e análises no banco de dados do National Center for Biotechnology Information (NCBI). Esta edição inclui três tutoriais novos: banco de dados COG, análise funcional e genômica ambiental.

Animações (totalmente em inglês)



Quarenta e cinco animações foram desenvolvidas por Anthony Griffiths e estão totalmente integradas ao conteúdo e às ilustrações nos capítulos. Disponíveis para serem vistas online em www.whfreeman.com/iga9e ou como vídeos para melhor transporte e flexibilidade. Os estudantes agora podem rever processos e conceitos genéticos importantes, de acordo com sua conveniência, fazendo download das animações em seus aparelhos de MP3.

CAPÍTULO 2

Mitose (Boxe 2.1)

Meiose (Boxe 2.2)

Estrutura Tridimensional de Cromossomos Nucleares (Figuras 2.4 e 2.5)

CAPÍTULO 3

Recombinação Meiótica entre Genes Não-ligados por Distribuição Independente (Figuras 3.8 e 3.13)

CAPÍTULO 4

Recombinação Meiótica entre Genes Ligados por Crossing Over (Figura 4.7)

Análise RFLP e Mapeamento Gênico (Figura 4.15)

CAPÍTULO 5

Conjugação Bacteriana e Mapeamento por Recombinação (Figuras 5.11 e 5.17)

CAPÍTULO 6

Interações entre Alelos a Nível Molecular, *RR*: Interações Tipo Selvagem entre Alelos a Nível Molecular, *rr*: Homozigoto Recessivo, Mutação Nula

Interações entre Alelos a Nível Molecular, *r'r'*: Homozigoto Recessivo, Mutação Leaky

Interações entre Alelos a Nível Molecular, *Rr*: Heterozigotos, Dominância Completa (Figura 6.4)

CAPÍTULO 7

Replicação do DNA: O Processo de Polimerização de Nucleotídeos (Figura 7.15)

Replicação do DNA: Coordenação da Síntese dos Filamentos de Síntese Contínua e de Síntese Descontínua (Figura 7.18)

Replicação do DNA: Replicação de um Cromossomo (Figura 7.22)

CAPÍTULO 8

Transcrição (Figura 8.4)

CAPÍTULO 9

Tradução: Formação de Ligação Peptídica (Figura 9.2)

Tradução: As Três Etapas da Tradução (Figura 9.17)

Supressão sem Sentido a Nível Molecular: O Supressor sem Sentido de tRNA (Figura 9.19)

Supressão sem Sentido a Nível Molecular: A Mutação sem Sentido *rod^{ns}* (Figura 9.19)

Supressão sem Sentido a Nível Molecular: Supressão sem Sentido do Alelo *rod^{ns}* (Figura 9.19)

CAPÍTULO 10

Regulação do Sistema Lactose em *E. coli*: Avaliando a Presença ou Ausência de Lactose Através do Repressor Lac (Figura 10.6)

Regulação do Sistema Lactose em *E. coli*: Mutações no Sistema Operador *O^c* Lac (Figura 10.8)

Regulação do Sistema Lactose em *E. coli*: Mutações de Repressor *I⁻* lac (Figura 10.9)

Regulação do Sistema Lactose em *E. coli*: Mutações de Super-repressor *I^s* lac (Figura 10.10)

CAPÍTULO 12

Interações Gênicas na Embriogênese de *Drosophila*: Sinalização Celular na Formação do Eixo Dorsoventral

CAPÍTULO 13

Microarranjos de DNA: Uso de um Arranjo de Oligonucleotídeos para Analisar Padrões de Expressão Gênica (Figura 13.19)

Microarranjos de DNA: Síntese de um Arranjo de Oligonucleotídeos

CAPÍTULO 14

Transposição Replicativa (Figura 14.10)

CAPÍTULO 15

Mecanismo Molecular de Mutação (Figura 15.10)

Fotodímeros Induzidos por UV e Reparo por Excisão (Figura 15.22)

Um Mecanismo de Crossing Over: Modelo Heterodúplex de Meselson-Radding

Um Mecanismo de Crossing Over: Consequências Genéticas do Modelo de Meselson-Radding

CAPÍTULO 16

Meiose Autotetraplóide (Figura 16.6)

Não-disjunção Meiótica na Meiose I (Figura 16.12)

Não-disjunção Meiótica na Meiose II (Figura 16.12)

Rearranjos Cromossômicos: Inversão Paracêntrica, Formação de Inversões Paracêntricas (Figura 16.27)

Rearranjos Cromossômicos: Inversão Paracêntrica, Comportamento Meiótico de Inversões Paracêntricas (Figura 16.28)

Rearranjos Cromossômicos: Translocação Recíproca, Formação de Translocações Recíprocas (Figura 16.30)

Rearranjos Cromossômicos: Translocações Recíprocas, Comportamento Meiótico das Translocações Recíprocas (Figura 16.30)

Rearranjos Cromossômicos: Translocações Recíprocas, Pseudoligação de Genes por Translocações Recíprocas (Figura 16.32)

CAPÍTULO 20

Reação em Cadeia da Polimerase (Figura 20.14)

Encontrando Genes Clonados Específicos por Complementação Funcional: Complementação Funcional da Linhagem de Levedura Gal⁻ e Recuperação do Gene *GAL* Tipo Selvagem

Encontrando Genes Clonados Específicos por Complementação Funcional: Fazendo uma Biblioteca de DNA de Levedura Tipo Selvagem

Encontrando Genes Clonados Específicos por Complementação Funcional: Uso do Gene Clonado *GAL* como Sonda para mRNA de *GAL*

Triagem e Seleção de Mutações

Agradecimentos

Tornamos extensivos nossa gratidão e nosso reconhecimento a todos os colegas que revisaram esta edição e cujos *insights* e conselhos foram muito úteis:

Paul Babitzke
Pennsylvania State University

Miriam Barlow
University of California, Merced

J. T. Beatty
University of British Columbia

Craig Berezowsky
University of British Columbia

Andrew J. Bohonak
San Diego State University

Robb T. Brumfield
Louisiana State University

Michael A. Buratovich
Spring Arbor University

Joan Burnside
University of Delaware

Kimberly A. Carlson
University of Nebraska, Kearny

J. Aaron Cassill
University of Texas, San Antonio

Yury O. Chernoff
Georgia Institute of Technology

Gregory P. Copenhaver
*University of North Carolina,
Chapel Hill*

Johnny El-Rady
University of South Florida

Dr. Robert Farrell
Pennsylvania State University, York

Wayne Forrester
Indiana University, Bloomington

Rick Gaber
Northwestern University

Maria Gallo
University of Florida

T. J. Gill
University of Houston

Dr. Michael A. Goldman
San Francisco State University

Jody L. Hall
Brown University

Pamela L. Hanratty
Indiana University, Bloomington

Margaret Hollingsworth
*State University of New York,
Buffalo*

John B. Jenkins
Swarthmore College

Jeffrey M. Marcus
Western Kentucky University

Kim McKim
Rutgers University

Philip Meneely
Haverford College

Steve Mount
*University of Maryland,
College Park*

Dr. David G. Muir
Stanford University

Bryan Ness
Pacific Union College

Dr. David K. Peyton
Morehead State University

Dr. Deborah Polayes
George Mason University

Michael Polymenis
Texas A & M University

James V. Price
Utah Valley State College

David Rivier
*University of Illinois,
Urbana-Champaign*

Daniel Rokhsar
University of California, Berkeley

Megan E. Rokop
Massachusetts Institute of Technology

Charles E. Rozek
Dean of Graduate Studies
Case Western Reserve University

Christine Rushlow
New York University

James J. Russo
Columbia University, Genome Center

Inder Saxena
University of Texas, Austin

Malcolm D. Schug
*University of North Carolina,
Greensboro*

Mark A. Seeger
Ohio State University

J. Kenneth Shull, Jr.
Appalachian State University

Marla B. Sokolowski
University of Toronto

Douglas Thrower
*University of California,
Santa Barbara*

Fyodor Urnov
University of California, Berkeley

Daniel Wells
University of Houston

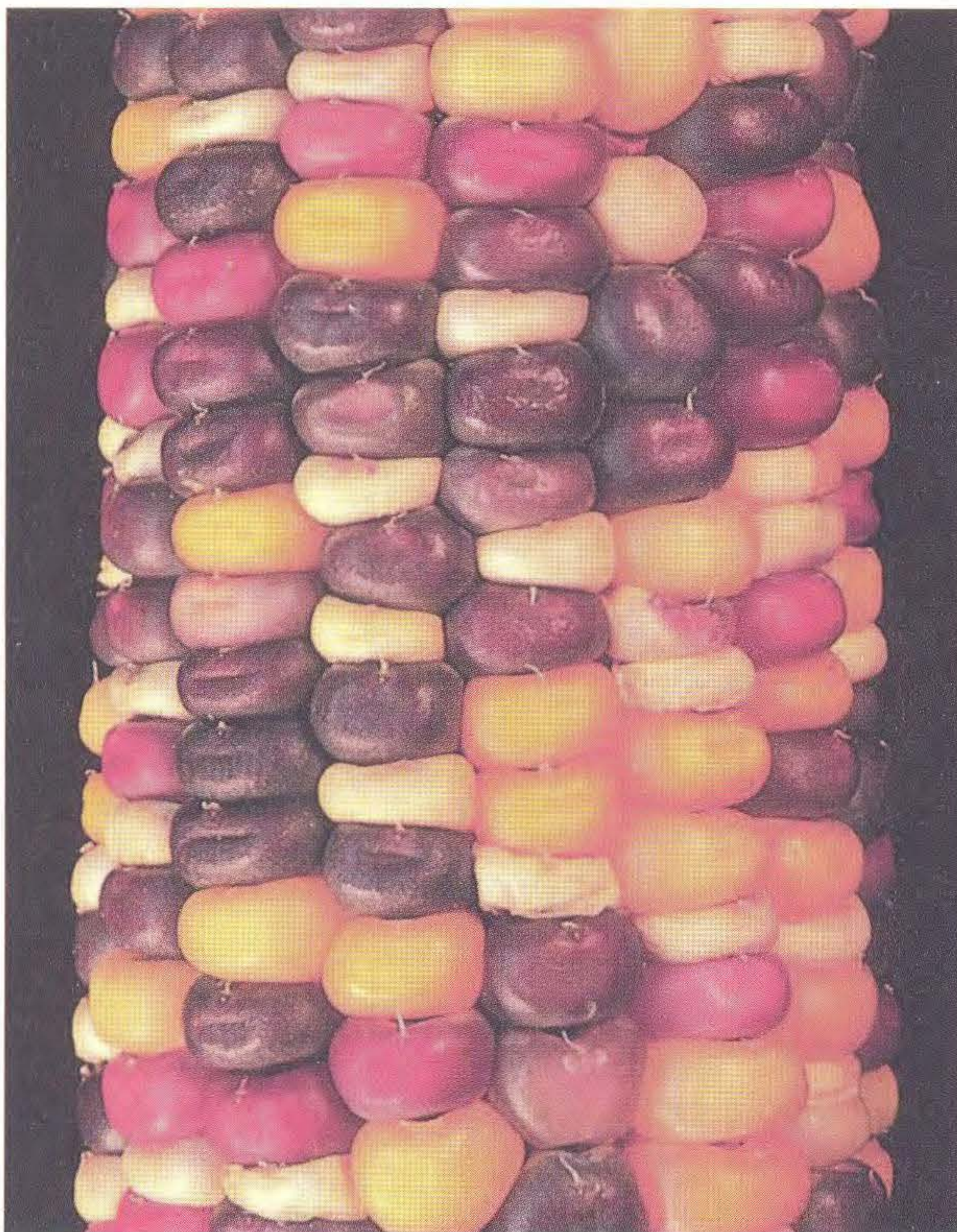
Patricia J. Wittkopp
University of Michigan

Os autores também agradecem à equipe da W.H. Freeman por seu trabalho árduo e paciente. Em particular, agradecemos à nossa editora de desenvolvimento, Susan Moran, que, com sua “mão de ferro em luva de veludo”, fez muitas grandes contribuições para a estrutura e conteúdo do livro. O editor de aquisição Jerry Correa fez um grande trabalho, sempre nos mantendo no foco e na trilha. Somos gratos também ao bom trabalho de Patty Zimmerman e Mary Louise Byrd, editoras cujos cuidados e atenção a detalhes foram cruciais para preparar um texto complexo, escrito por quatro autores, sobre um assunto eminentemente técnico. Bill Fixsen deu uma grande contribuição à utilidade pedagógica dos conjuntos de problemas no livro por meio de seu caprichoso manual de soluções selecionadas e respostas. Também agradecemos a Paul Rohloff, coordenador de produção; Diana Blume, coordenadora de design; Sheridan Sellers, que fez um ótimo trabalho nos layouts das páginas; Bill Page, coordenador das ilustrações; Ted Szczepanski, editor de fotos; e Janie Chan, assistente editorial. Agradecemos em especial os esforços de marketing e vendas de Debbie Clare, diretora de marketing, e de toda a equipe de vendas. Amy Peltier, editora de mídia, e Lindsay Lovier, editora de suplementos, habilmente reuniram o amplo pacote de mídia e suplementos.

Introdução à Genética

O Enfoque Genético à Biologia

1



Variação genética na cor dos grãos de milho. Cada grão representa um indivíduo separado com uma constituição genética distinta. A foto simboliza a história do interesse da humanidade pela hereditariedade. Os humanos cruzaram milho por milhares de anos antes do advento da disciplina da Genética. Ampliando essa herança, o milho hoje é um dos principais organismos de pesquisa na Genética Clássica e Moderna. [William Sheridan, University of North Dakota; foto de Travis Amos.]

É uma observação diária de que a prole de camundongos são camundongos e a prole de humanos são humanos. O que é extraordinário sobre esse fato é que cada camundongo e cada ser humano começa a vida como uma única célula, o resultado da fertilização de um ovócito por um espermatozóide, e que os ovócitos fertilizados de camundongos e humanos são muito parecidos. Eles certamente não têm semelhança com os organismos adultos aos quais darão origem. Todos os organismos multicelulares, plantas e animais, sofrem um processo de desenvolvimento que transforma o zigoto na planta ou animal adulto. As novas células surgem como resultado de grandes números de divisões celulares, e essas células diferenciam-se tanto morfológica quanto bioquimicamente para atender a uma variedade de funções, tornando-se, por exemplo, células sanguíneas, células nervosas, células musculares e células secretoras. Por meio dessa diferenciação, as células tornam-se organizadas em

Perguntas Fundamentais

- O que é o material genético?
- Qual é a estrutura química e física do DNA?
- Como o DNA é copiado na formação de novas células e nos gametas que darão origem à prole de um organismo individual?
- Quais são as unidades funcionais do DNA que levam informações sobre o desenvolvimento e a fisiologia?
- Que moléculas são os principais determinantes das propriedades básicas estruturais e fisiológicas de um organismo?
- Quais são as etapas na tradução da informação no DNA em proteína? Quais são as causas da variação entre os membros individuais de uma espécie?

Tópicos

- 1.1 Genética e as perguntas da biologia
- 1.2 A base molecular da informação genética
- 1.3 O programa da investigação genética
- 1.4 Metodologias usadas em genética
- 1.5 Organismos-modelo
- 1.6 Genes, o ambiente e o organismo

tecidos e órgãos, cujas formas e funções são características de cada espécie. Assim, as propriedades biológicas de um organismo são o produto final de uma sequência de etapas de desenvolvimento.

Isso significa que o zigoto contém a informação que especifica a sequência de etapas de desenvolvimento e as propriedades biológicas do organismo que são o produto final desse desenvolvimento. Essa informação foi supostamente passada de genitor para a prole nos gametas que se juntaram para formar o zigoto. Qualquer que seja a base física da passagem da informação, ela deve ter quatro propriedades:

1. *Diversidade de estrutura.* As estruturas celulares que levam a informação sobre o desenvolvimento e a função devem ser capazes de existir em um imenso número de formas diferentes, cada uma especificando um aspecto diferente do organismo complexo. Deve haver informação sobre o desenvolvimento de todos os tipos diferentes de células, tecidos, órgãos e processos bioquímicos que caracterizam a espécie.
2. *Capacidade de se replicar.* A fusão de um gameta masculino e um gameta feminino produz um único camundongo, embora esse camundongo venha, na maturidade sexual, a produzir um número imenso de gametas que contêm a informação necessária para fazer outra geração de camundongos. Deve haver algum mecanismo de replicação para copiar a estrutura portadora de informação, de modo que ela possa ser passada do genitor para a prole.
3. *Mutabilidade.* Todos os membros individuais de uma espécie não são idênticos; mesmo assim, a prole tende a se assemelhar aos genitores. As diferenças na cor da pele e forma do cabelo que diferenciam típicos africanos sub-saarianos dos típicos europeus do Norte são herdadas por seus filhos, não importando em que continente eles nasceram. Contudo, os europeus do Norte são descendentes de grupos que saíram da África há dezenas de milhares de anos e na época dessa migração assemelhavam-se a seus ancestrais africanos. Nos anos intercalares, as mudanças nas informações sobre a cor da pele e forma dos cabelos ocorreram e se espalharam pela prole tanto dos emigrantes quanto dos parentes que ficaram em seus lares. As estruturas portadoras das informações, então, têm de ser capazes de sofrer mudanças, chamadas **mutações**.

Todas as espécies que existem surgiram da evolução de espécies ancestrais que se diferenciaram delas em uma variedade de características. Camundongos e humanos têm um ancestral em comum. Assim, em alguma época do passado, devem ter surgido mutações que alteraram algumas das informações passadas entre os pais e a prole, e essas mudanças devem ter sido herdadas. Tais mutações podem resultar em alterações relativamente pequenas em alguma propriedade de um organismo, tal como cor da pele, ou em grandes alterações na forma, tais como as que diferenciam os humanos dos camundongos. O acúmulo de um grande número de mutações pode até mesmo resultar na formação de estruturas totalmente novas, tais como os membros dos vertebrados.

4. *Tradução.* Não é suficiente que exista um mecanismo apenas para transferir informações sobre um organismo de uma geração para outra por meio do zigoto. Um mapa deve conter a informação necessária para especificar um automóvel, por exemplo; mas, sem a maquinaria da fábrica, funcionários e energia, não seria feito o carro. Analogamente, o zigoto e as células que surgem de suas divisões devem ter uma maquinaria que possa ler a informação herdada e traduzi-la em uma grande diversidade de estruturas biológicas que constituem a forma e a função do organismo. Além disso, esses eventos de tradução devem ocorrer em épocas determinadas do desenvolvimento do organismo e em algumas partes do organismo, mas não em outras. A informação sobre a estrutura química da hemoglobina, por exemplo, deve ser traduzida em moléculas de hemoglobina começando em certo estágio do desenvolvimento de um feto, e esse evento traducional deve ocorrer na medula óssea, mas não nas células neurais que irão constituir o cérebro.

Os elementos básicos do sistema de informação herdada hoje em dia são chamados **genes**, um termo introduzido em 1909 por Wilhelm Johannsen, que investigava a herança em feijões. A coleção de todos os genes em um organismo é chamada de **genoma**. O campo da **Genética** envolve a diversidade, replicação, mutação e tradução da informação nos genes.

► 1.1 Genética e as Perguntas da Biologia

A ciência da Genética começou com o trabalho do monge austríaco Gregor Mendel, que publicou o resultado de seus experimentos em cruzamentos entre linhagens que haviam herdado variações em ervilhas em 1865. Mendel (Figura 1.1) não só deu os resultados experimentais de cruzamentos controlados, como também deduziu a existência de distintos “fatores” que levavam a informação sobre o desenvolvimento dos genitores para a prole. No começo do século XX, tornou-se aparente que a informação que especifica o desenvolvimento dos organismos estava contida nos cromossomos do núcleo celular.

Os padrões de herança de características nos cruzamentos entre linhagens de diferentes variantes herdadas, descobertos e descritos sistematicamente por Mendel, correspondem aos padrões de distribuição dos cromossomos nos gametas na meiose. À medida que as pesquisas sobre herança continuavam, foram descobertas informações sobre várias características situadas em pequenos trechos de material, os *genes*, situados em pontos diferentes ao longo de cromossomos. A estrutura química do material genético e o modo pelo qual os genes especificam as informações permaneceram desconhecidos até a metade do século XX. A despeito dessa falta de conhecimento, foi possível uma grande quantidade de manipulação genética no período inicial da ciência da genética. Os cientistas logo perceberam que essas manipulações genéticas deram um meio para investigar propriedades biológicas em

O fundador da Genética



FIG. 1.1 Gregor Mendel. [Moravian Museum, Brno.]

geral. Como os genes afetam praticamente toda a estrutura e função de um organismo, ser capaz de identificar e determinar o papel dos genes é uma etapa importante ao mapear os vários processos subjacentes a determinada propriedade particular. É interessante que os geneticistas estudem não apenas os mecanismos hereditários, mas *todos* os mecanismos biológicos.

A chave para a análise genética de uma propriedade biológica é examinar os efeitos das mutações. Foi descoberto que bombardear um organismo com raios X ou tratá-lo com substâncias químicas induz um grande número de mutações com efeitos claros sobre a morfologia do organismo. Algumas mutações correspondem a deleções observáveis ou duplicações de pequenos trechos do material cromossômico. Pelos resultados dos cruzamentos entre as linhagens portadoras de mutações diferentes, algumas perguntas importantes podem ser respondidas. Duas alterações de mutações ocorrem no mesmo gene ou em genes diferentes? Quais as consequências desenvolvimentais e fisiológicas de combinar mutações diferentes no mesmo organismo?

Essas técnicas experimentais de induzir mutações, que as localizam em cromossomos e as combinam de diferentes modos, continuam a fornecer um poderoso método para investigar tanto o desenvolvimento quanto as vias fisiológicas e bioquímicas que constituem a biologia dos organismos. Um

exemplo dos estudos da mosca-das-frutas *Drosophila* ilustra o que pode ser aprendido dessas técnicas. Nos anos de 1920 e 1930, foram descobertas algumas mutações na mosca-das-frutas que fazem com que antenas se desenvolvam em estruturas tipo pernas, enquanto outras mutações e combinações de mutações fazem aparecer um par extra de asas, transformando uma mosca de duas asas em um inseto com quatro asas. Outros estudos de mutantes desses tipos revelaram muito sobre como os membros se desenvolvem em todos os organismos com vários membros. Além disso, essas mutações deram um forte suporte para a posição dos evolucionistas e anatomistas de que as asas, pernas e antenas, em organismos diferentes, são variações de uma estrutura básica, podendo ser convertidas de uma para outra na evolução.

O acúmulo de evidências que começou nos anos de 1920 levou à conclusão de que o DNA é o material genético. O poder analítico da investigação genética aumentou imensamente mais recentemente pelo conhecimento da estrutura química do DNA, de como a informação no DNA é convertida em moléculas de atividade celular e como sinais bioquímicos das células controlam que informação genética será lida por células em partes diferentes de um organismo em épocas diferentes em sua história de vida. A Figura 1.2 resume a conexão entre o organismo todo, sua constituição celular e as estruturas subjacentes cromossômicas e moleculares que levam a informação sobre o organismo.

É importante perceber que não há estágio na vida de um organismo em que a informação genética pare de ser lida. As moléculas com atividade fisiológica estão sendo produzidas o tempo todo, em parte em resposta a mudanças na fisiologia do organismo e, em parte, porque as moléculas no organismo estão sendo degradadas e precisam ser substituídas. Além disso, as células têm um tempo de vida limitado, e novas devem ser produzidas por divisão celular com a produção de novas cópias de seus conteúdos. Portanto, a replicação do DNA, a mutação e o processamento da informação especificada pelo DNA continuam ao longo da vida do organismo. Assim, a compreensão de fenômenos genéticos é um enfoque unificador para a compreensão de uma grande variedade de processos na vida de um organismo em todos os estágios, não apenas durante o desenvolvimento de um adulto a partir de um zigoto.

A análise genética, usando organismos com mutações, é um método poderoso para investigar vias bioquímicas, fisiológicas e desenvolvimentais.

A genética também é a base de nossa compreensão da evolução. A mudança evolutiva é a consequência das diferentes taxas de reprodução de organismos tendo características herdáveis diferentes. Como resultado dessa reprodução diferencial de diferenças herdáveis, a população sofre mudanças na frequência na qual variantes diferentes são representadas em gerações sucessivas. A população evolui à medida que surgem novas variantes herdadas e tornam-se mais comuns. O processo de evolução orgânica não pode ser compreendido sem invocar vários fenômenos genéticos, tais como mutação,

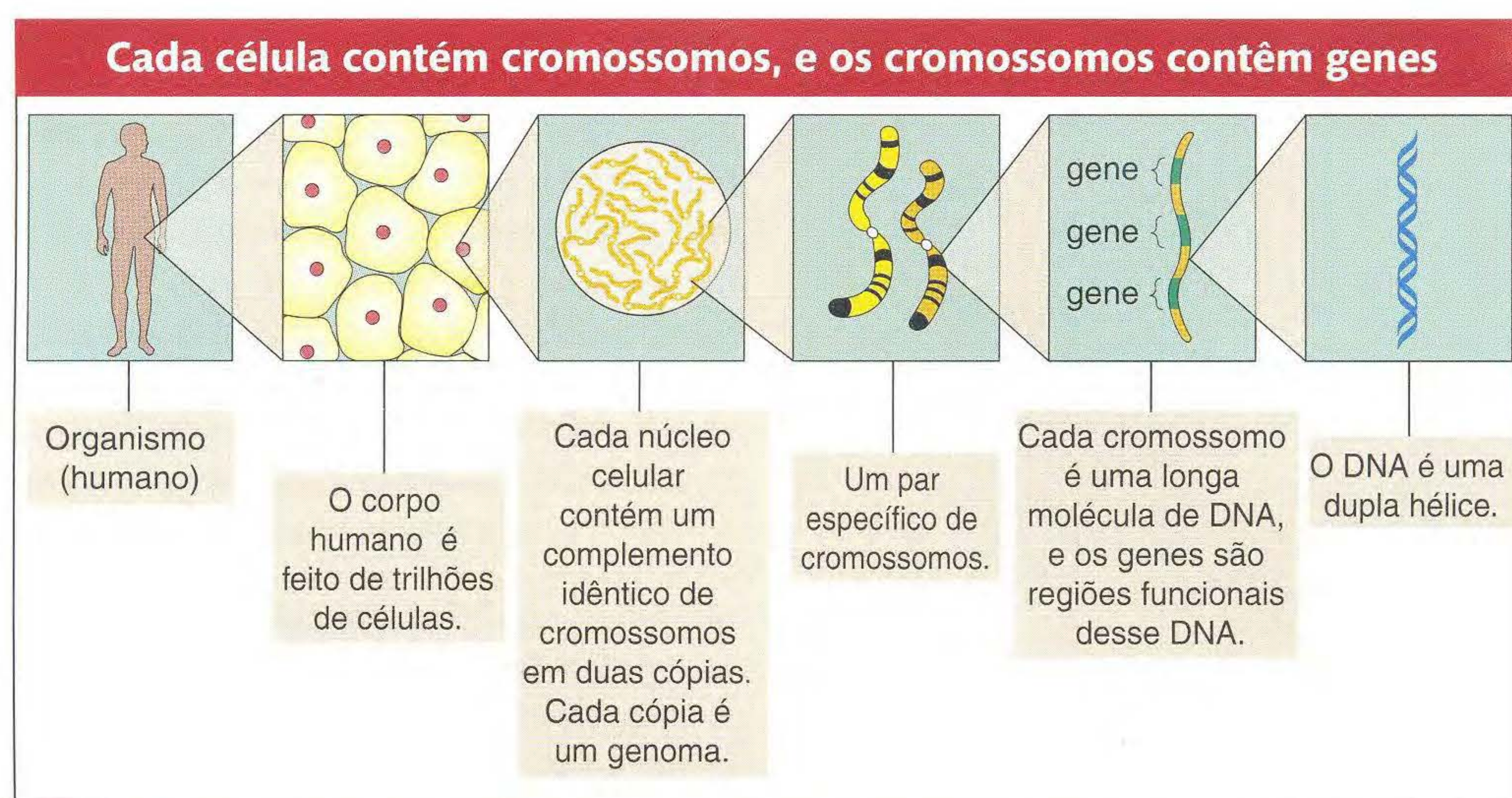


FIG. 1.2 Ampliações sucessivas levam o material genético de um organismo a um foco mais nítido.

padrões de reprodução e a tradução da informação genética na forma e funcionamento de um indivíduo. As explicações evolutivas são, em grande parte, explicações genéticas.

No passado, os biólogos evolutivos reconstruíram correlações evolutivas entre espécies diferentes comparando um número limitado de diferenças morfológicas e fisiológicas entre elas. Não era possível saber o quanto dois organismos tinham ficado diferentes em todas as propriedades biológicas desde sua divergência evolutiva. Com o advento das técnicas de seqüenciamento de DNA em larga escala, os biólogos evolutivos agora podem comparar genomas inteiros de espécies, bem como determinar a quantidade de divergências que ocorreu em diferentes partes desses genomas. Viu-se, por exemplo, que ocorreu muito pouca mudança genética na divergência dos humanos e chimpanzés em sua evolução desde seu ancestral primata comum. Todavia, essas diferenças estão manifestas no sistema nervoso central, na musculatura da língua e lábios, no controle motor fino dos apêndices superiores e na postura, mudanças que fazem uma imensa diferença entre a vida de humanos e de chimpanzés.

Vemos então que, cada vez mais, as técnicas genéticas de análise forneceram possibilidades inigualáveis para a compreensão de problemas biológicos. A genética não é apenas um dos muitos aspectos do estudo da biologia. A análise genética é uma via de enfoque para quase todas as propriedades dos sistemas vivos.

► 1.2 A Base Molecular da Informação Genética

A estrutura dos organismos e seus ativos processos fisiológicos são baseados, na maior parte, em **proteínas**. A informação genética para a síntese dessas proteínas pelas células está contida no **DNA, ácido desoxirribonucleico**. Uma molécula de DNA é feita de dois filamentos enrolados um no outro em uma longa dupla hélice. Cada um dos dois filamentos consiste em um arcabouço feito de cópias repetidas de um

açúcar, chamado desoxirribose, e fosfato; de cada grupo açúcar-fosfato ao longo do arcabouço, projeta-se uma **base nucleotídica**. Existem quatro tipos diferentes de bases nos nucleotídeos no DNA: **adenina (A), timina (T), guanina (G) e citosina (C)**. Na molécula em dupla hélice, o arcabouço açúcar-fosfato de cada filamento está por fora da hélice, enquanto cada base nucleotídica projeta-se para dentro e faz par com o filamento oposto (Figura 1.3). Devido ao espaço ocupado pelas bases nucleotídicas, a adenina sempre parecia com timina, enquanto a guanina sempre parecia com citosina. As bases que formam pares são ditas **complementares**.

O DNA é composto de duas cadeias nucleotídicas mantidas unidas pelo pareamento de A com T e G com C.

Essa estrutura molecular fornece a base para as quatro propriedades que caracterizam a informação genética.

Diversidade de estrutura Embora existam apenas quatro tipos de nucleotídeos em um único filamento de DNA, esses nucleotídeos podem estar em qualquer ordem, e o trecho de DNA correspondente a determinado gene pode ter qualquer tamanho. Assim, um tipo de gene pode ter a seqüência ...TTACG-GACCT... nas posições 147 a 156, enquanto outro gene pode ter ...GCATACGATC... nas mesmas posições. Mesmo que os genes tenham apenas 100 bases de tamanho (eles são muito maiores), haveria apenas 4^{100} , ou cerca de $1.000.000^{10}$, tipos diferentes possíveis. Como ocorre, cada seqüência diferente não necessariamente leva informação diferente, mas o número de seqüências efetivamente diferentes ainda é enorme.

Habilidade de se replicar Devido ao pareamento de A em um filamento com T no outro e C em um filamento com G no outro, cada filamento contém uma especificação completa de seu filamento pareado complementar. Na replicação da dupla hélice, a primeira etapa é a separação dos dois filamentos

O DNA tem a forma de uma dupla hélice

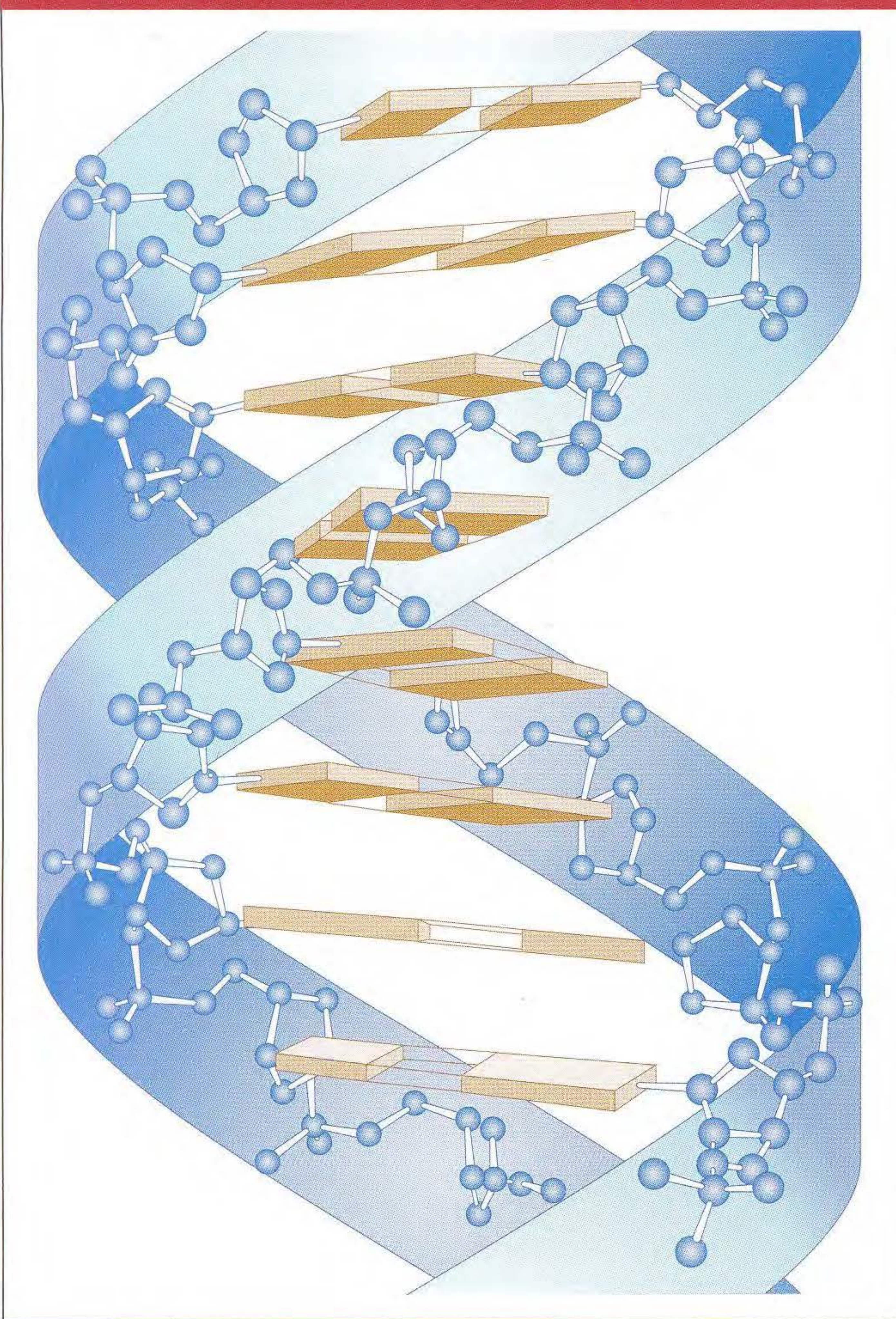


FIG. 1.3 Fita representando a dupla hélice de DNA. As bases pareadas (marrom) são ligadas a um arcabouço helicoidizado açúcar-fosfato (azul).

complementares. Então é feito um novo filamento-filho em cada um dos filamentos separados. Onde o filamento-molde tem um A, o novo filamento feito contém um T, um T fará par com A, um C com G e um G com C. O resultado é que são produzidas duas cópias idênticas da dupla hélice original (Figura 1.4).

O DNA é replicado pela deslicoidização dos dois filamentos da dupla hélice e pela construção de um novo filamento complementar sobre cada um dos filamentos separados da dupla hélice original.

Mutabilidade No curso da replicação, base incorreta pode ser colocada ou bases podem ser perdidas ou duplicadas. Caso ocorra tal evento, a nova cópia de DNA e todas as cópias decorrentes dessa cópia serão diferentes da molécula ancestral. Terá ocorrido uma mutação herdável.

Tradução em forma e função De algum modo, determinada sequência de A, T, G e C deve ser usada pela célula para criar moléculas de proteínas com sequências particulares de aminoácidos. Além disso, alguma parte do DNA deve agir como um sinal para a maquinaria da célula de que a tradução de determinado gene em uma sequência de aminoácidos deve ocorrer em algumas células, em determinados tecidos e em certas épocas do desenvolvimento e vida do organismo. Como a maquinaria da célula usa a sequência de bases para determinar tanto as sequências de aminoácidos nas proteínas quanto a época e o local de produção dessas moléculas proteicas?

Especificando a sequência de aminoácidos de uma proteína

A transformação da informação na sequência de nucleotídeos em uma sequência de aminoácidos em uma molécula

Cada filamento de DNA é o molde para a produção de um novo filamento

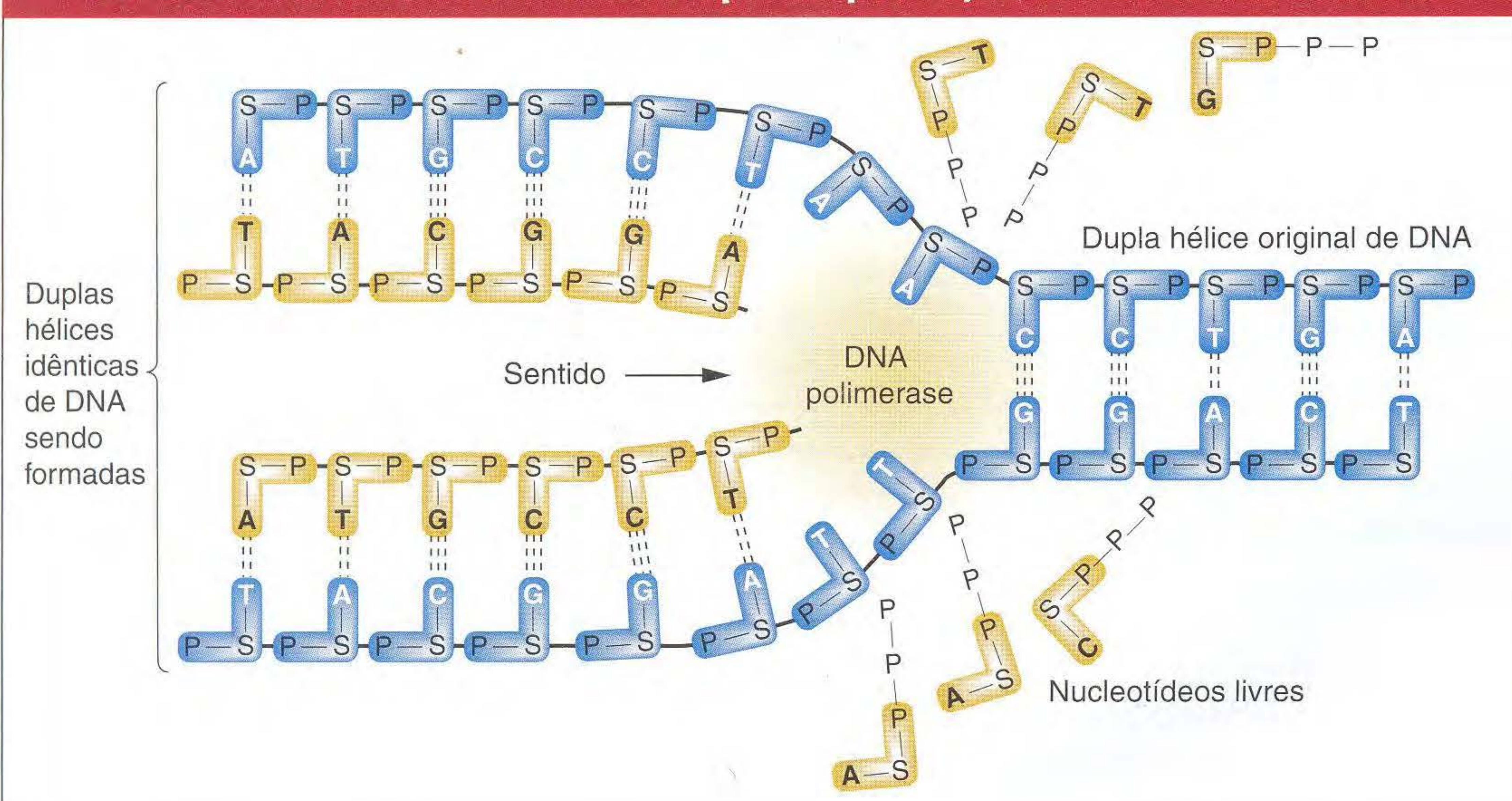


FIG. 1.4 Na replicação do DNA, novos nucleotídeos (amarelo) são polimerizados para formar cadeias-filhas, usando nucleotídeos da dupla hélice original (azul) como molde. Abreviações: S, açúcar; P, grupo fosfato.

Transcrição e tradução em um eucarionte

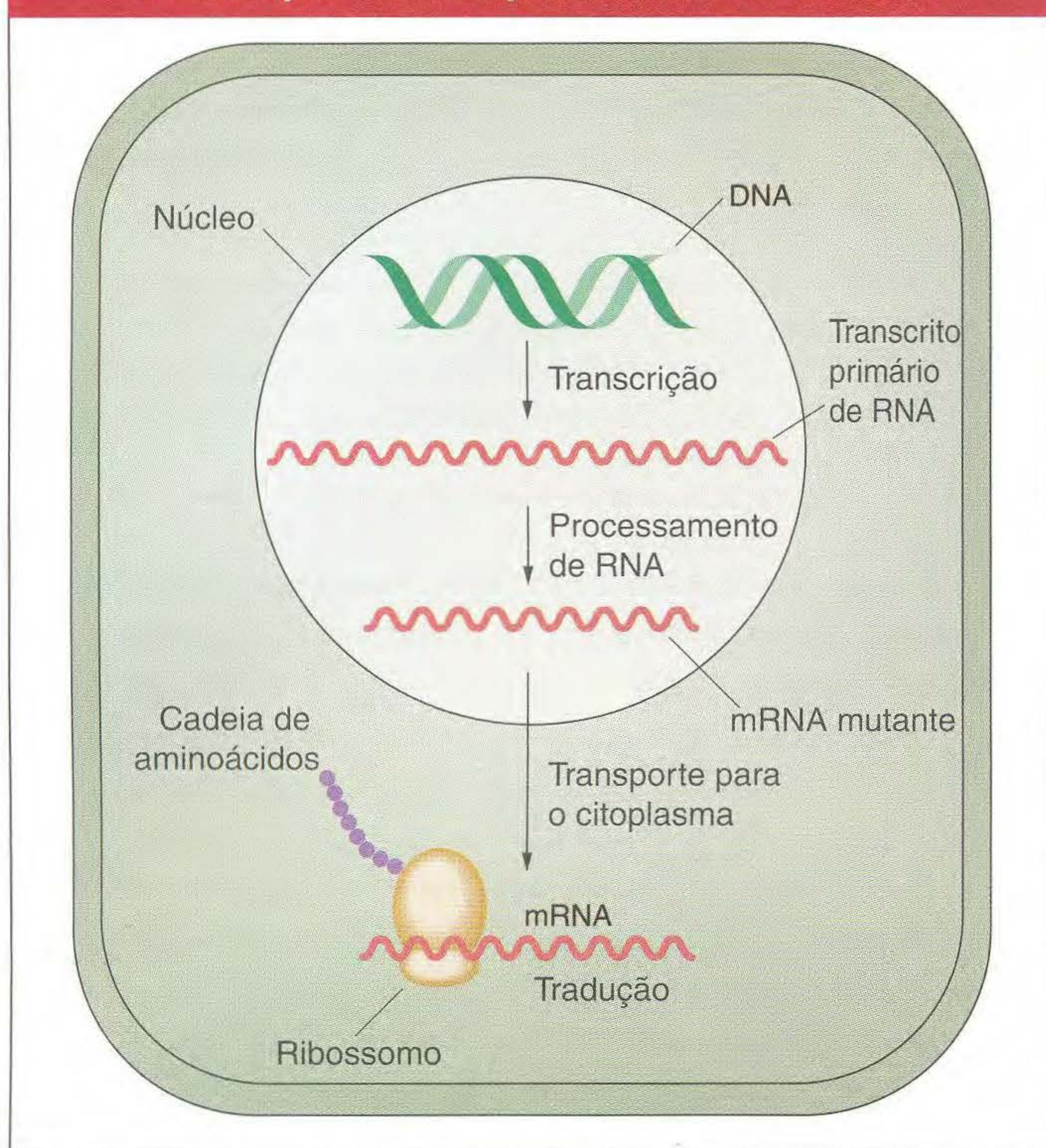


FIG. 1.5 Em uma célula eucariótica, o mRNA é transcrito do DNA no núcleo e, então, transportado para o citoplasma para tradução em uma cadeia polipeptídica.

que irá dobrar-se para fazer uma proteína compreende duas etapas: **transcrição** e **tradução** (Figura 1.5).

Transcrição A primeira etapa em transformar a informação no DNA em uma sequência de aminoácidos é a transcrição da sequência de nucleotídeos no DNA em uma molécula correlata, o **RNA mensageiro (mRNA)**, que também tem um arcabouço de nucleotídeos como estrutura. O RNA mensageiro é composto da molécula de **ácido ribonucleico (RNA)**, que é uma sequência de nucleotídeos similar ao DNA exceto por conter ribose em vez de desoxirribose em seu arcabouço e ter o nucleotídeo uracil (U) em lugar de timina. Na transcrição, a dupla hélice de DNA é separada em dois filamentos únicos, e um desses filamentos serve de molde para a construção de uma sequência de RNA unifilamentar. O processo de transcrição, que ocorre no núcleo, é assim muito similar ao processo de replicação do DNA porque o filamento de DNA serve como molde para fazer um filamento de mRNA. O mRNA cópia da molécula original de DNA é chamado de **transcrito** (veja Figura 1.5). O RNA que é produzido diretamente da transcrição do DNA pode então ser alterado para produzir um mRNA final (RNA mensageiro). Essa alteração consiste em remover trechos do transcrito original que não codificam aminoácidos.

O RNA final transcrito é uma espécie de “cópia funcional” do DNA. A produção desses transcritos serve a três funções importantes na célula. Primeira, aumenta o número de cópias da informação genética disponível para a célula a qualquer

tempo. Embora a célula contenha apenas uma cópia de cada uma das moléculas de DNA que é herdada de cada um de seus genitores, a transcrição pode produzir um grande número de cópias funcionais de cada gene. Além disso, cada uma dessas cópias funcionais pode ser usada mais de uma vez pela maquinaria funcional em sua produção de proteínas. Segunda, o RNA mensageiro resolve o problema da congestão do tráfego. Como seu nome indica, o RNA mensageiro deixa a vizinhança imediata dos genes na célula onde muitos genes estão sendo transcritos e geralmente passa para o citoplasma da célula, onde está disponível à maquinaria celular de produção de proteínas. Terceira, a estabilidade e o tempo de vida das moléculas mensageiras atuam como controles do quanto de determinada proteína será produzido.

Durante a transcrição, um dos filamentos de DNA de um gene serve como molde para a síntese de uma molécula complementar de RNA.

Tradução A produção de uma cadeia de aminoácidos com base em uma sequência de nucleotídeos no processo de tradução. Como uma sequência feita apenas de quatro tipos diferentes de nucleotídeos pode gerar uma sequência de aminoácidos composta de 20 tipos diferentes de aminoácidos? A solução é que a sequência de nucleotídeos no mRNA é “lida” em grupos sucessivos de três nucleotídeos ao longo da cadeia de mRNA. Cada grupo de três é um **códon**.

AUU	CCG	UAC	GUA	AAU	UUG
Códon	Códon	Códon	Códon	Códon	Códon

Como existem então $4 \times 4 \times 4 = 64$ trincas códon e apenas 20 aminoácidos, mais de um códon pode corresponder a cada aminoácido. Por exemplo, AUC, AUU e AUA, todos codificam o aminoácido leucina. Existem também códons de “fim”, tais como UAG, que indicam o final da sequência traduzida.

As moléculas que de fato contêm em forma física a correspondência entre os códons e os aminoácidos são pequenas moléculas dobradas de RNA, chamadas de **RNA transportador (tRNA)**. Em uma ponta de uma molécula de tRNA para determinado aminoácido, há um **anticódon** feito de três nucleotídeos que são o complemento do códon para esse aminoácido e, assim, irão parear com o códon apropriado no mRNA. Na outra ponta de cada tRNA está o aminoácido que está sendo codificado. Um tRNA com seu aminoácido ligado é dito *carregado*. O que falta fazer então é colocar o aminoácido carregado por diferentes tRNA na ordem especificada pela mensagem. Essa etapa ocorre em um componente da maquinaria de síntese, um **ribossomo**.

A Figura 1.6 mostra esquematicamente como a cadeia polipeptídica de aminoácidos é feita. Um ribossomo passa por uma molécula de RNA mensageiro, e a cada códon, quando o tRNA complementar pareia com o códon, a maquinaria ribossômica transfere o aminoácido no tRNA para a cadeia polipeptídica crescente. O ribossomo então se move para o códon seguinte no mRNA, repetindo o processo de transferência até atingir o códon de “fim”, quando ele se desliga do

O RNA mensageiro é traduzido em uma cadeia polipeptídica

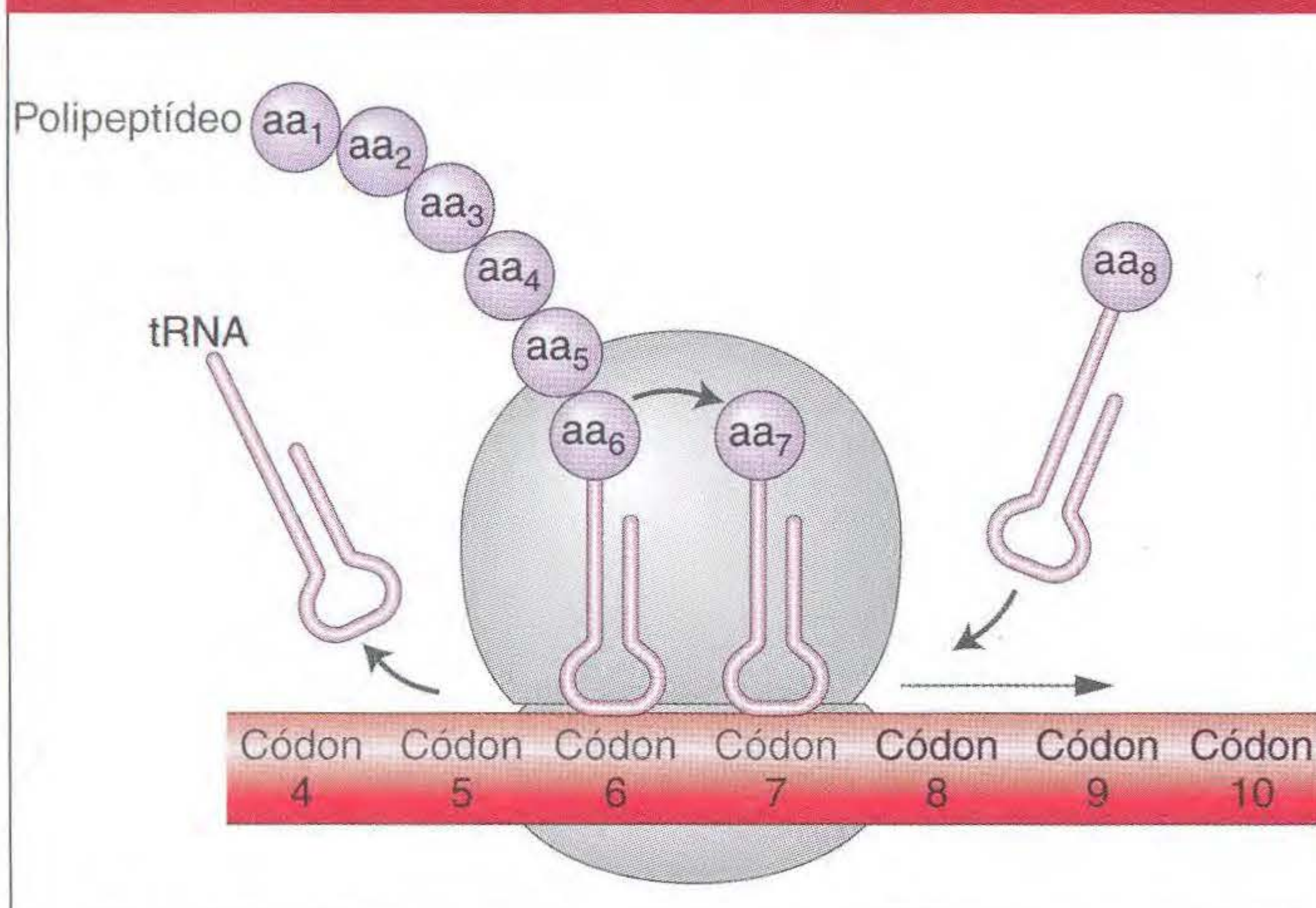


FIG. 1.6 Um aminoácido (aa) é adicionado ao polipeptídeo crescente na tradução do mRNA.

mRNA e está pronto para ser usado novamente. Uma molécula de tRNA que acabou de ser usada e não tem mais seu aminoácido ligado é então recarregada na célula com o aminoácido apropriado de modo a ser usada novamente. Existem muitos ribossomos movendo-se ao longo de uma molécula de mRNA e um grande número de moléculas de tRNA de cada tipo, e, assim, o processo de construir o polipeptídeo move-se muito rapidamente. Uma explicação mais detalhada desse processo é dada nos Capítulos 8 e 9.

A informação nos genes é usada pela célula em duas etapas de transferência de informação: o DNA é transcrito em mRNA, que é então traduzido na sequência de aminoácidos de um polipeptídeo. O fluxo de informação do DNA para o RNA e para a proteína é o foco central da biologia moderna.

Do polipeptídeo à proteína É importante compreender que a estrutura primária que resulta da síntese, a cadeia polipeptídica, não é a molécula que nós chamamos de uma “proteína”. Uma proteína é uma cadeia polipeptídica dobrada. Determinada sequência de aminoácidos, embora confine o processo de dobramento, não o determina totalmente. De fato, para cada determinada sequência de aminoácidos, existem vários dobramentos estáveis alternativos. Além disso, a célula pode fazer algumas alterações químicas nos aminoácidos da molécula. O dobramento final que produz uma proteína biologicamente ativa depende das condições na célula que incluem a presença de uma variedade de outras moléculas. Essa dependência tem efeitos práticos. A insulina humana, usada pelos diabéticos, hoje é produzida em culturas de células para as quais o gene de insulina humana foi transferido. Quando primeiro produzida desse modo, a “insulina” não tem atividade fisiológica, muito embora tenha a sequência correta de aminoácidos. A insulina fisiologicamente ativa foi depois produzida por uma mudança no método de produção que resultou em uma molécula modificada.

A sequência de nucleotídeos em um gene especifica a sequência de aminoácidos colocados juntos pela célula para produzir um polipeptídeo. Esse polipeptídeo, então, dobra-se sob a influência de sua sequência de aminoácidos e outras condições moleculares na célula para formar uma proteína.

Regulação gênica

Para construir um organismo a partir de um zigoto, deve haver também um mecanismo para regular em que momentos e em que locais no organismo a informação no genoma deve ser traduzida. A maquinaria molecular dessa regulação também é feita pelo DNA. De fato, a maioria do DNA codifica essa função reguladora.

Como já vimos, a leitura de um gene pela célula envolve a transcrição dos genes do DNA em RNA mensageiro. Nesse processo, a longa molécula de mRNA é sintetizada etapa por etapa a partir de ribonucleotídeos individuais que correspondem aos nucleotídeos na sequência de DNA. Essa síntese sequencial requer uma enzima, a RNA polimerase, que deve encontrar o começo da sequência de DNA e, então, continuar ao longo do gene ligada a ribonucleotídeos sucessivos.

A regulação gênica ocorre tanto na ligação da RNA polimerase ao começo da sequência de DNA que está para ser lida quanto no início de sua movimentação ao longo dessa sequência de DNA. As moléculas dentro da célula podem se ligar a uma região controladora do DNA no começo de um gene e bloquear o movimento da RNA polimerase ao longo da sequência do gene, impedindo assim a transcrição (controle negativo). Alternativamente, uma molécula que normalmente se liga ao DNA e impede a transcrição pode ser deslocada do DNA por alguma outra molécula dentro da célula, de modo que a transcrição agora pode continuar (controle positivo). Em ambos os casos, a informação do estado fisiológico da célula controla se o gene é lido. Esse controle de informação depende da estrutura e fisiologia da célula, de onde a célula está situada no organismo e de que sinais estão vindo do ambiente. Por exemplo, as bactérias que obtêm sua energia do açúcar lactose fazem uma enzima que quebra a lactose para fornecer essa energia. Entretanto, a enzima β -galactosidase, que degrada lactose, é feita pela bactéria apenas quando lactose está presente na célula. Na ausência de lactose, a transcrição do gene para a enzima é bloqueada por uma molécula repressora que se situa no DNA no sítio do gene. Quando as moléculas de lactose estão presentes na célula, elas combinam-se com o repressor e o removem do DNA, permitindo que a transcrição continue. Os detalhes desse processo de controle são discutidos nos Capítulos 10 e 11.

► 1.3 O Programa da Investigação Genética

A necessidade da variação

A essência do método genético como meio para explicar fenômenos biológicos é o uso das diferenças genéticas entre os

organismos. Dentro das espécies, existem diferenças de ocorrência natural entre membros individuais das populações. Por exemplo, cerca de um terço de todos os genes codificantes de proteínas em espécies de reprodução sexual apresenta alguma variação entre indivíduos nas seqüências de aminoácidos que eles codificam. Esse **polimorfismo genético** é uma fonte rica de material para o estudo das bases da variação biológica entre indivíduos.

Quando nenhuma variação herdada de ocorrência natural está disponível para determinada característica, a variação genética pode ser produzida no laboratório usando radiação de alta energia ou substâncias químicas para produzir mutações. Entretanto, sendo obtida, uma variante genética torna-se uma “incisão” no processo de dissecação genética. Já vimos que o genoma geral fornece uma fonte de informações para o desenvolvimento de uma planta, animal ou espécie microbiana. Do mesmo modo que um corpo pode ser cirurgicamente dissecado usando um bisturi para tentar compreender suas funções, uma variação genética pode ser usada para separar as vias do desenvolvimento normal do organismo. A dissecação genética em geral é feita em um gene de cada vez: cada gene é identificado e sua função é investigada. Após os efeitos individuais das mutações que agem em uma via particular de desenvolvimento ou bioquímica serem observados e analisados, eles são montados para dar uma visão geral do processo geneticamente especificado.

Em adição à compreensão da base das diferenças entre os membros individuais de uma espécie, pode ser usado o método genético para estudar os eventos evolucionários que resultaram na grande diversidade de forma e função entre as formas vivas. As diferenças na forma e função entre as espécies, o resultado da divergência evolutiva, podem estar relacionados a diferenças nos genomas desses organismos pela técnica de seqüenciamento de DNA. No momento, cerca de uma dúzia de genomas de espécies multicelulares foram sequenciadas, incluindo humanos e chimpanzés, bem como várias espécies da mosca-das-frutas *Drosophila*, que é um dos grupos de organismos experimentais mais estudados geneticamente. Além disso, os genomas de algumas centenas de espécies de bactérias foram seqüenciados.

A variação genética é a diferença genética entre os indivíduos. A variação genética é o foco dos métodos genéticos de estudo das propriedades biológicas.

Começando com variação: Genética direta

Uma investigação genética pode começar com a observação de uma variação detectada na morfologia ou fisiologia, isto é, um **fenótipo** variante. Assim, a primeira etapa na **genética direta** é pesquisar a diferença genética que causa uma diferença fenotípica. Tal pesquisa envolve procurar os padrões de herança nos descendentes de cruzamentos entre indivíduos que têm fenótipos diferentes. Quando é identificado um gene, um programa de investigação é então desenvolvido para mostrar como a informação contida nesse gene é nor-



FIG. 1.7 O fenótipo albino é causado por duas doses de um gene variante incomum (*alelo*) *a*. A variante normal *A* determina uma etapa na síntese química do pigmento escuro melanina nas células da pele, cabelos e retina dos olhos. Nos indivíduos *a/a*, essa etapa não é funcional, e a síntese de melanina é bloqueada. [Copyright Yves Gellie/Icone.]

malmente transformada em eventos celulares e fisiológicos que contribuem para as características normais do organismo. Uma forma de tal investigação começa com a variação natural normal no fenótipo. Por exemplo, pestes de insetos individuais podem apresentar sensibilidades diferentes a algum inseticida. Nesse caso, um gene pode ser encontrado tendo uma variante diferente correspondente a cada variação em sensibilidade. Pode ser então que o gene especifique uma enzima que quebre o inseticida e que uma variante da enzima tenha uma afinidade maior pelo inseticida do que outra variante.

Outra forma de programa de genética direta começa com uma variante anormal em uma característica para a qual quase todos os indivíduos tenham uma forma normal. A forma normal é chamada de **tipo selvagem**, enquanto as raras variantes excepcionais são chamadas de **mutantes**. Por exemplo, a maioria dos seres humanos tem uma certa quantidade do pigmento melanina na pele e nos cabelos, mas algumas pessoas, chamadas albinas, não têm esse pigmento (Figura 1.7). A ausência de melanina é uma consequência de levar duas cópias de um gene mutante que codificam uma forma inativa de uma enzima que tem um papel na formação de melanina. O propósito primário de investigar um fenótipo mutante tal como o albinismo é elucidar a via normal do metabolismo e desenvolvimento. Um subproduto dessa investigação pode ser uma explicação de como a variante genética interferiu no processo normal, como no albinismo (Figura 1.8), mas essa explicação não é a meta da investigação.

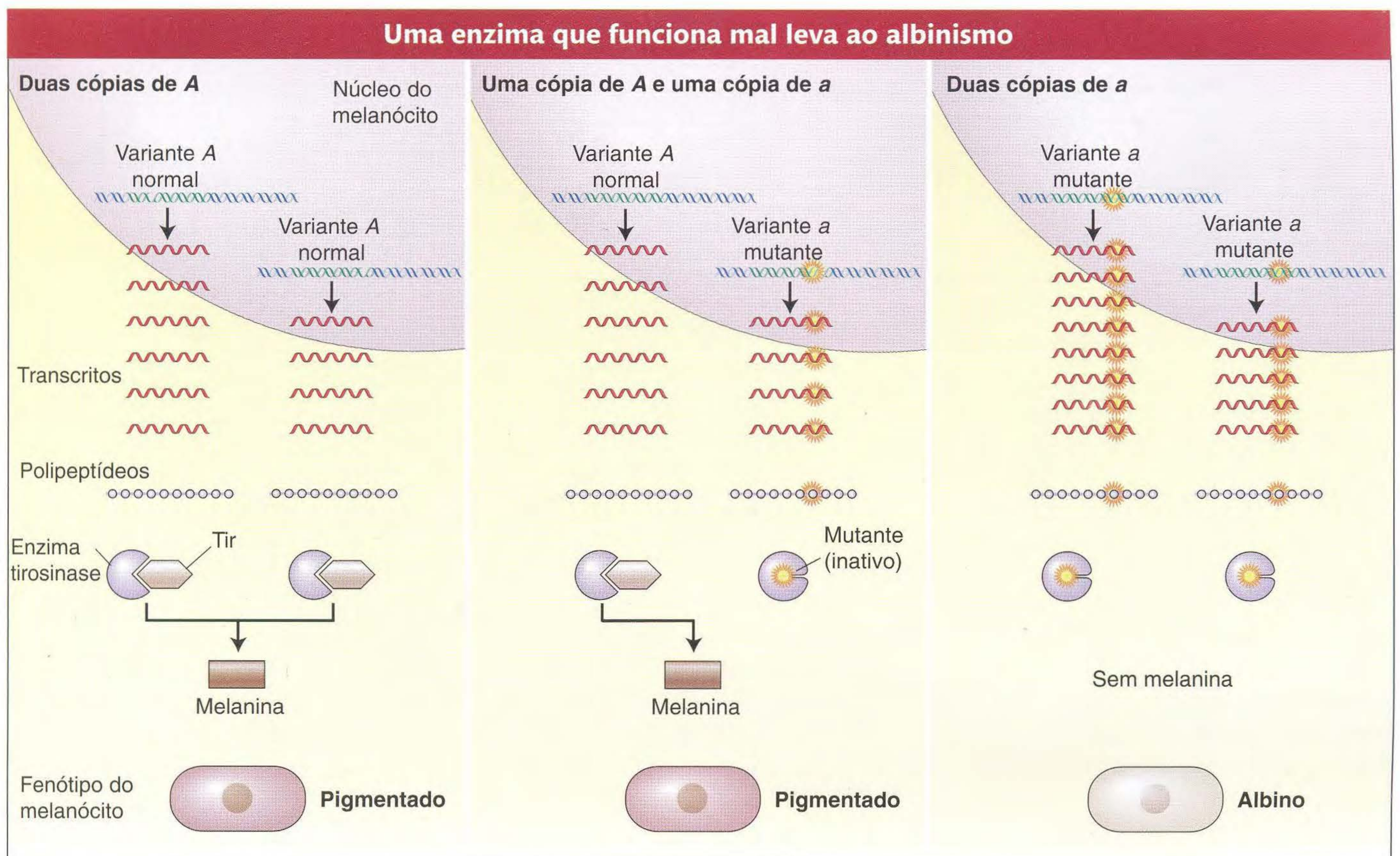


FIG. 1.8 Base molecular do albinismo. *Esquerda:* Melanócitos (células produtoras de pigmento) contendo duas cópias do gene normal tirosinase (A) produz a enzima tirosinase, que converte o aminoácido tirosina (Tir) no pigmento melanina. *Centro:* Melanócitos contendo uma cópia do gene normal fazem tirosinase suficiente para permitir a produção de melanina e o fenótipo pigmentado. *Direita:* Melanócitos contendo duas cópias do gene variante (a) são incapazes de produzir o que quer que seja da enzima.

A análise genética direta começa com indivíduos de dois fenótipos distintos. Um gene pode ser identificado de cruzamentos e uma análise dos padrões de herança na prole. A função do produto desse gene pode então ser investigada para esclarecer a via bioquímica, fisiológica ou desenvolvimental.

A maioria das características fenotípicas são afetadas por várias vias fisiológicas ou desenvolvimentais que interagem, cada via consistindo em um número de etapas. Tais casos requerem vários genes. Na formação da pele e pigmento dos cabelos, por exemplo, claramente devemos procurar a via de produção de melanina, mas onde na via? As enzimas operam em moléculas que tomam parte na produção de melanina (tais como o aminoácido triptofano, um intermediário), mas também existem processos celulares que controlam o depósito de grânulos de melanina na pele e cabelos. O método que foi usado é a pesquisa de várias mutações, sejam induzidas natural ou artificialmente, que afetam a característica que está sendo estudada. Essas mutações são então individualmente mapeadas em suas posições cromossômicas no genoma (como descrito no Capítulo 4). Fazer cruzamentos entre as linhagens coloca essas mutações diferentes juntas em combinações diferentes, permitindo que os pesquisadores determinem seus efeitos conjuntos no fenótipo, como dis-

cutido no Capítulo 6. Para cada mutação e combinação de mutações, são feitas observações a nível do organismo todo, tecido ou célula para procurar a presença ou ausência de várias enzimas e outras proteínas, sua atividade metabólica, sua distribuição espacial e padrões de movimento de células e tecidos. Quando essas observações são contrastadas com os resultados de observações similares em organismos normais, podemos obter um quadro das vias causais do metabolismo normal e desenvolvimental. Esse estágio da análise genética é o assunto dos Capítulos 6 e 12.

Um exemplo clássico é ilustrado na Figura 1.9. A mosca-das-frutas *Drosophila* normalmente tem olhos vermelho-escuros. Uma mutação gênica *vermilion* torna os olhos vermelho-claros, e uma mutação em um gene diferente, *brown*, torna os olhos marrons. Mas uma linhagem que tem ambas as mutações tem olhos brancos. Nós podemos deduzir que os olhos normais vermelhos têm uma mistura de pigmentos *vermilion* e marrom sintetizados em vias separadas, e que a mutação *vermilion* bloqueia a via produzindo pigmento marrom e a mutação *brown* bloqueia a via produzindo pigmento *vermilion*. Uma série de mutações diferentes de um outro gene, *white*, ou bloqueia toda a deposição de pigmento, produzindo olhos brancos, ou dilui a cor do olho de modo que ela é amarelo bem claro ou cor de palha. O gene normal no qual ocorre a mutação *white* está então funcionando adiante

Mutações que interagem afetam a cor do olho em *Drosophila*

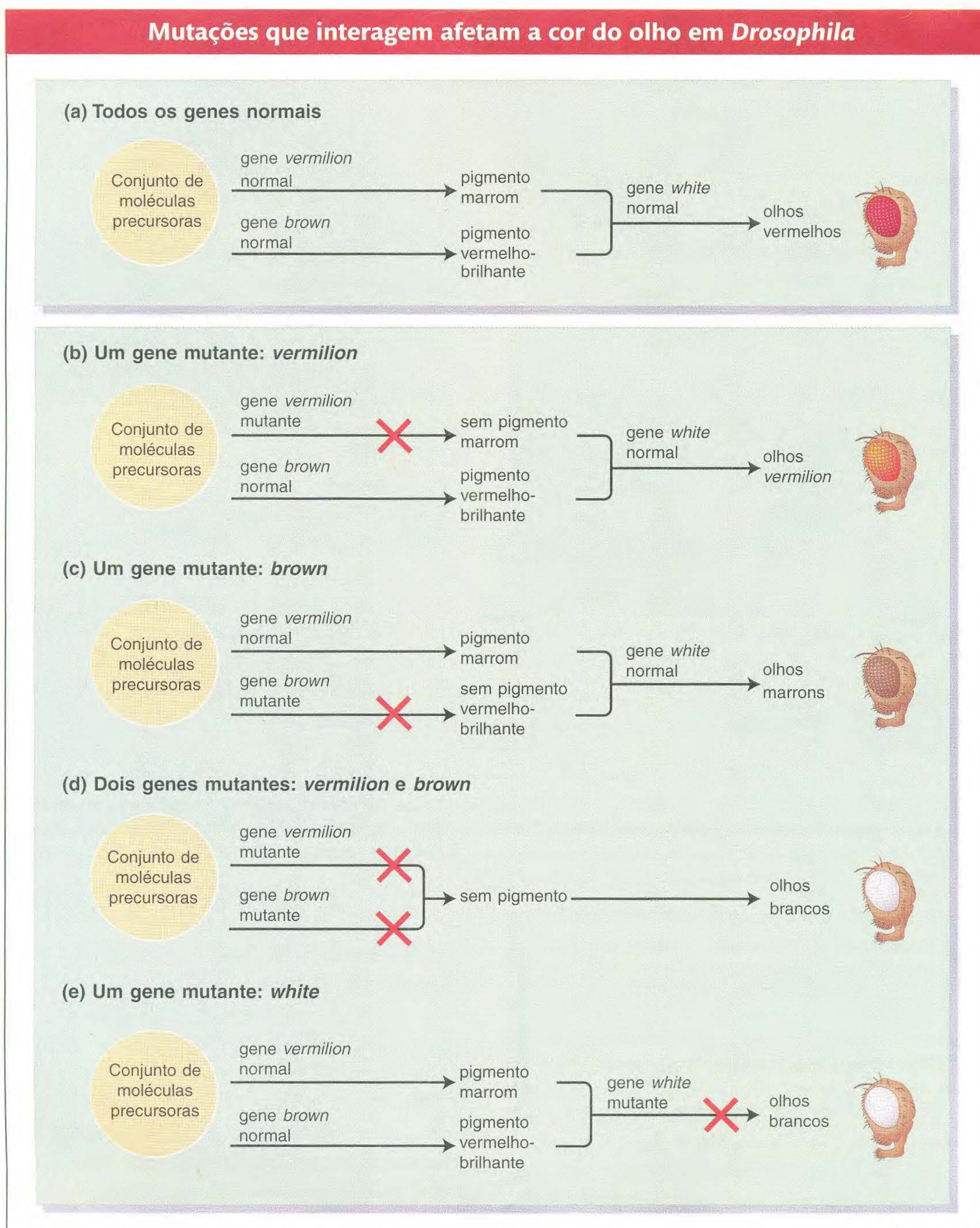


FIG. 1.9 (a) Uma via esquemática parcial para a produção do pigmento vermelho normal a partir de moléculas precursoras em *Drosophila*. Note que o pigmento normal é uma combinação de *vermilion* e marrom. (b-e) Os efeitos das mutações impedindo a produção desses pigmentos.

na via de formação de pigmento, após terem sido formados os pigmentos vermelho e marrom, como ilustrado esquematicamente na Figura 1.9.

Os produtos de vários genes são ativos nas vias que determinam propriedades biológicas tais como cor de olho. Para separar os efeitos de múltiplos genes, são usados cruzamentos para criar indivíduos com mutações em combinações diferentes de genes. Observando os efeitos, os pesquisadores podem começar a construir a via biológica que determina a propriedade.

O último estágio do programa direto é caracterizar o DNA de variantes diferentes (alelos) dos genes implicados e explicar, a partir dessas seqüências variantes de DNA, as diferenças em estrutura, função, quantidades e localização dentro do organismo das moléculas metabolicamente ativas. Algumas mudanças no DNA irão corresponder a diferenças na seqüência de aminoácidos das proteínas. Essas diferenças de aminoácidos podem responder totalmente por mudanças na atividade biológica. Outras mudanças de DNA não alte-

A análise genética direta começa com indivíduos de dois genótipos diferentes

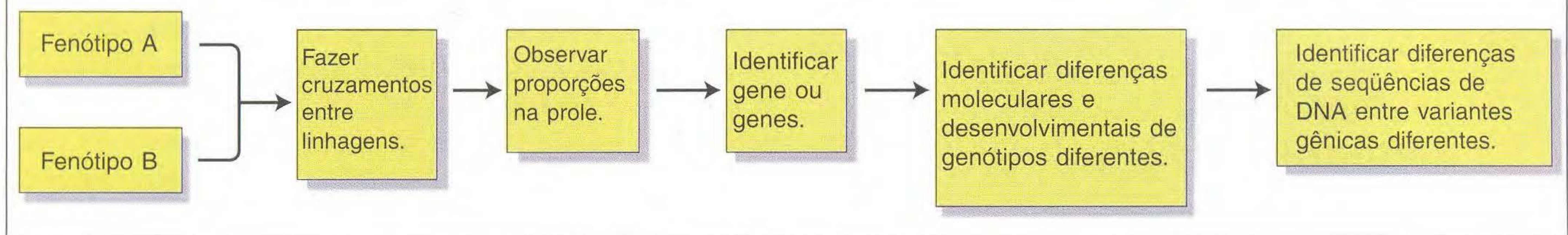


FIG. 1.10 Um fluxograma do programa da genética direta.

rarão a proteína em si, mas a taxa de sua produção ou as condições celulares que disparam sua síntese.

Um fluxograma mostrando as etapas em um programa genético direto de análise é mostrado na Figura 1.10.

Começando com o DNA: Genética reversa

Com o desenvolvimento de nossos conhecimentos de DNA e o modo pelo qual ele codifica a informação sobre a seqüência de aminoácidos nas proteínas e o controle celular da produção dessas proteínas, tornou-se possível uma forma alternativa de investigação genética. Vimos que o programa de genética direta começa com as diferenças observadas na forma e função do organismo e procura as diferenças genéticas causais. Em contraste, o programa da **genética reversa** começa com mudanças genéticas conhecidas e procura as alterações resultantes no organismo.

Começando com uma seqüência normal de DNA, podemos usar sua informação para ler a seqüência de aminoácidos na proteína que é produzida, pois sabemos a correspondência entre o código do DNA e os aminoácidos que são codificados. O DNA pode então ser alterado de modos específicos que, ou bloqueiam totalmente a produção da proteína, ou mudam a

proteína de tal modo que ela fica com a atividade metabólica gravemente alterada ou sem atividade metabólica. O efeito de tais mutações “nocaute” no desenvolvimento do fenótipo podem então ser seguidos do mesmo modo que os efeitos mutacionais no programa genético direto. Uma vantagem do programa reverso é que pode ser ajustado um grande número de tipos específicos de mutações.

Outra forma de programa reverso não requer criar mutantes dentro de uma espécie. Ele faz uso de informações já disponíveis sobre como os genomas de uma variedade de espécies relacionadas diferem uns dos outros. Está obviamente fora de questão deliberadamente alterar os genes de um ser humano para compreender, digamos, como o genoma humano faz com que o cérebro humano seja muito maior em proporção ao peso do corpo do que em outros primatas. Podemos, entretanto, comparar o genoma humano com o dos chimpanzés, nossos parentes mais próximos entre os primatas. Agora que os genomas, tanto do *Homo sapiens* quanto de *Pan troglodytes*, foram seqüenciados e descoberto que diferem muito pouco um do outro, temos alguma esperança em encontrar uma explicação para coisas tais como diferenças no tamanho dos cérebros (Figura 1.11).

Os genomas de humanos e chimpanzés diferem apenas em uma pequena percentagem de nucleotídeos

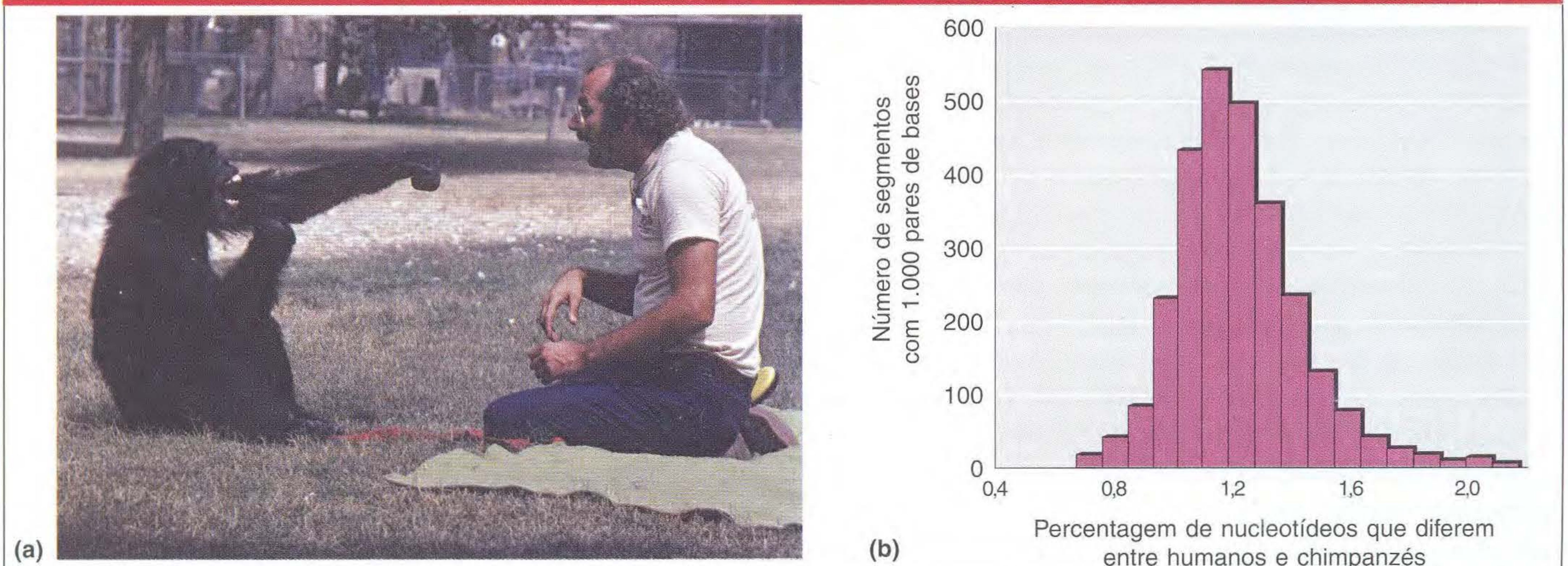


FIG. 1.11 (a) Esses dois primatas diferem um do outro apenas em 1% do DNA. (b) A comparação das seqüências de nucleotídeos de humanos e chimpanzés. Um gráfico de barras mostra o número de segmentos do genoma, de 24.000 segmentos diferentes examinados (eixo y), que diferem entre o chimpanzé e o humano por uma percentagem de todos os nucleotídeos em um segmento (eixo x). Cada segmento examinado tinha 1.000 nucleotídeos de tamanho. A diferença média era de cerca de 1,2% por segmento. [(a) Vic Cox/Peter Arnold.]

Uma análise genética reversa começa com uma sequência normal de DNA. Inserindo uma mutação no DNA (ou comparando-a com o DNA de outros genomas), podemos analisar a função da sequência de DNA.

► 1.4 Metodologias Usadas em Genética

Visão geral

Já vimos que o estudo dos genes demonstrou ser um poderoso enfoque para compreender os sistemas biológicos. Muitas metodologias diferentes são usadas para estudar genes e atividades gênicas, e essas metodologias podem ser resumidas do seguinte modo:

1. *Isolamento de mutações que afetam o processo biológico em estudo.* Cada gene mutante revela um componente genético do processo, e juntos os genes mutantes mostram a gama de proteínas que interagem nesse processo.
2. *Análise da prole de cruzamentos controlados entre mutantes e indivíduos tipo selvagem ou outras variantes descontínuas.* Esse tipo de análise identifica genes e seus alelos, suas localizações cromossômicas e seus padrões de herança. Esses métodos serão introduzidos no Capítulo 2.
3. *Análise genética dos processos bioquímicos da célula.* A vida é basicamente um conjunto complexo de reações químicas, e, assim, estudar os modos pelos quais os genes são relevantes para essas reações é um modo importante de dissecar essa química complexa. Os alelos mutantes subjacentes a uma função defeituosa (veja método 1) são valiosos nesse tipo de análise. O enfoque básico é descobrir como a química celular é perturbada no indivíduo mutante e, com essa informação, deduzir o papel do gene. As deduções das análises de muitos genes são reunidas para revelar um panorama maior.
4. *Análise microscópica.* A estrutura cromossômica e o movimento há muito têm sido uma parte integral da Genética, mas novas tecnologias forneceram meios de marcação dos genes e produtos gênicos que têm suas localizações facilmente visualizadas ao microscópio.
5. *Análise direta do DNA.* Como o material genético é composto de DNA, a caracterização final de um gene é a análise da própria sequência do DNA. Muitas técnicas, incluindo a clonagem gênica, são usadas para fazer essa análise. A clonagem é um procedimento pelo qual um gene individual pode ser isolado e amplificado (copiado muitas vezes) para produzir uma amostra simples para análise. Um modo de fazer isso é inserindo o gene de interesse em um pequeno cromossomo bacteriano e deixando que a bactéria faça o trabalho de copiar o DNA inserido. Após a clonagem de um gene ter sido obtida, sua sequência de nucleotídeos pode ser determinada, e assim podem ser obtidas importantes informações sobre sua estrutura e funcionamento.

Genomas inteiros de muitos organismos foram seqüenciados por extensões das técnicas anteriores, dando origem assim a uma nova disciplina dentro da Genética chamada **Genômica**, o estudo da estrutura, função e evolução de genomas inteiros. A **Genômica Comparativa** descreve as diferenças e similaridades nos genomas de espécies que são relacionadas por graus diferentes — por exemplo, duas espécies diferentes de primatas (veja Figura 1.11b). Parte da Genômica é a **Bioinformática**, a análise computacional do conteúdo de informação dos genomas.

Deteção de moléculas específicas de DNA, RNA e proteínas

Estudando genes individualmente ou genomas, os geneticistas sempre precisam detectar a presença de uma molécula específica de DNA, RNA ou proteína, as principais macromoléculas da genética. Essas técnicas serão totalmente descritas no Capítulo 20, mas precisamos de uma resumida visão geral delas, que podem ser usadas em capítulos anteriores.

Como moléculas específicas podem ser identificadas entre milhares de tipos na célula? O método mais amplo usado para detectar macromoléculas específicas em uma mistura é o de **sondas**. Esse método usa a especificidade da ligação intermolecular — por exemplo, a ligação de um mRNA com o DNA do qual foi transcrito. Uma mistura de macromoléculas é exposta a uma molécula, a sonda, que irá ligar-se apenas à macromolécula procurada. A sonda é marcada de algum modo, seja com um átomo radioativo ou com um composto fluorescente, de modo que o sítio de ligação possa ser facilmente detectado. Vejamos as sondas para DNA, RNA e proteínas (Figura 1.12).

Sondas para um DNA específico Um gene clonado pode agir como uma sonda para encontrar segmentos de DNA que têm a mesma sequência ou uma sequência muito similar. Por exemplo, se um gene de um fungo foi clonado, pode ser interessante determinar se as plantas têm o mesmo gene. O uso de um gene clonado como uma sonda nos leva de volta ao princípio da complementariedade das bases. A sonda funciona através do princípio de que, na solução, o movimento aleatório das moléculas da sonda permite-lhes encontrar sequências complementares e ligar-se a elas. O experimento deve ser feito com filamentos separados de DNA, porque assim os sítios de ligação das bases estão desocupados. O DNA da planta é extraído e cortado com um dos muitos tipos de **enzimas de restrição** disponíveis, que cortam o DNA em sequências-alvo específicas de quatro ou mais bases. As sequências-alvo estão nas mesmas posições em todas as células das plantas usadas, e assim a enzima corta o genoma em populações definidas de segmentos de tamanhos específicos. Os fragmentos podem ser separados em grupos de fragmentos do mesmo tamanho (fracionados) com o uso de eletroforese.

A eletroforese fraciona a população de fragmentos de ácido nucleico com base no tamanho. A mistura cortada é colocada em um pequeno poço em um gel, e o gel é colocado em um poderoso campo elétrico. A eletricidade faz com que

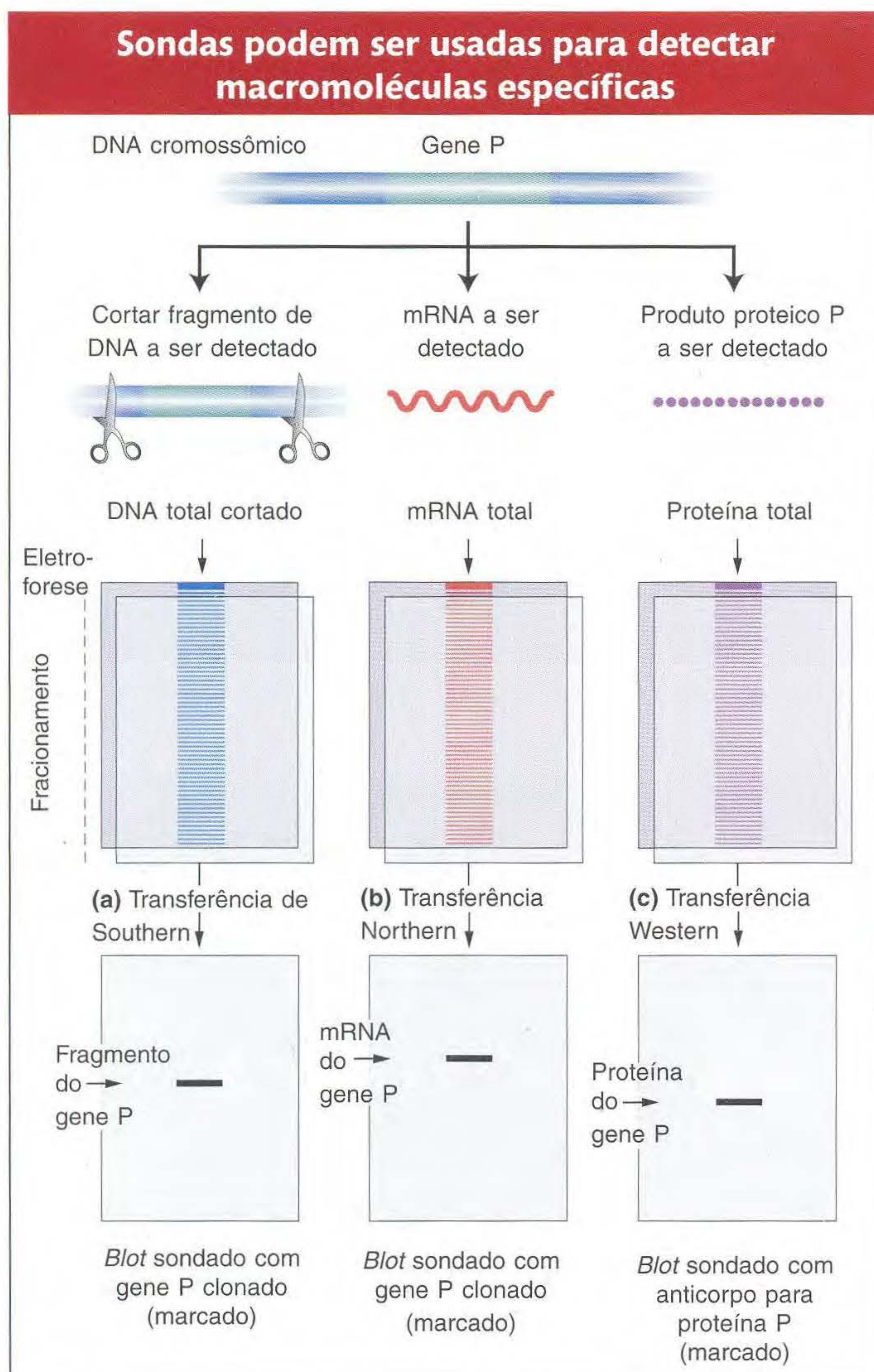


FIG. 1.12 Um gene específico pode ser usado como sonda para detectar o gene ou seu mRNA em uma mistura de DNA ou RNA, enquanto o anticorpo pode ser usado como sonda para detectar uma proteína específica em uma mistura de proteínas.

as moléculas se movam pelo gel a velocidades inversamente proporcionais a seu tamanho. Após o fracionamento, os fragmentos separados são transferidos para um pedaço de uma membrana porosa, onde eles mantêm as mesmas posições relativas. Esse procedimento é chamado de **transferência de Southern**. Após aquecer para separar os filamentos de DNA e manter o DNA posicionado, a membrana é colocada em uma solução da sonda. A sonda unifilamentar irá encontrar a sua sequência complementar de DNA e ligar-se a ela. Por exemplo,

TAGGTATCG	Sonda
ACTAATCCATAGCTTA	Fragmento genômico

Na transferência, essa ligação concentra a marca em um ponto, como mostrado na Figura 1.12a.

Sondagem de um RNA específico Em geral, é necessário determinar se um gene está sendo transcrito em algum tecido

particular. Para esse fim, é útil uma modificação da análise de Southern. O mRNA total é extraído do tecido, fracionado eletroforéticamente e transferido para a membrana (esse processo é chamado de **transferência Northern**). O gene clonado é usado como uma sonda, e sua marcação destacará o mRNA em questão se ele estiver presente (Figura 1.12b).

Sondagem de uma proteína específica A sondagem de proteínas geralmente é feita com anticorpos porque um anticorpo tem um ajuste específico tipo chave e fechadura com sua proteína-alvo ou antígeno. A mistura da proteína é separada em bandas de proteínas distintas por eletroforese e transferida para uma membrana (esse procedimento é uma **transferência Western**). A posição de uma proteína específica na membrana é revelada banhando a membrana em uma solução de anticorpo, obtida de um coelho ou outro hospedeiro no qual a proteína foi injetada. A posição da proteína é revelada pela posição da marca que o anticorpo leva (Figura 1.12c).

O DNA de um gene pode ser usado como uma sonda para encontrar segmentos similares em uma mistura de moléculas de DNA (transferência de Southern) ou de moléculas de RNA (transferência Northern). Pode ser usado um anticorpo como uma sonda para encontrar uma proteína específica em uma mistura de proteínas (transferência Western).

► 1.5 Organismos-modelo

A ciência da Genética discutida neste livro fornece uma compreensão das características da herança e desenvolvimento que são características dos organismos em geral. Algumas delas, especialmente a nível molecular, são verdadeiras para todas as formas de vida. Para outras, há uma variação entre grandes grupos de organismos — por exemplo, entre bactérias e todas as espécies multicelulares. Mesmo para características que variam, entretanto, essa variação está sempre entre os grandes grupos de formas vivas, de modo que não temos que investigar os fenômenos básicos da genética várias vezes para cada espécie. De fato, todos os fenômenos da Genética foram investigados por experimentos em um pequeno número de espécies, os **organismos-modelo**, cujos mecanismos genéticos são comuns ou para todas as espécies, ou para um grande grupo de organismos correlatos.

Lições dos primeiros organismos-modelo

O uso de organismos-modelo vem do trabalho de Gregor Mendel, que usou cruzamentos entre variedades hortícolas da ervilha de jardim, *Pisum sativum*, para estabelecer as regras básicas da herança. O uso feito por Mendel dessas variedades de ervilhas em estabelecer as “leis” básicas da herança é instrutivo para nossa compreensão das forças e fraquezas do estudo de organismos-modelo.

As leis de Mendel da herança foram a fundação da Genética e, em particular, estabeleceram que o mecanismo de herança era baseado em partículas distintas nos gametas, que vieram juntas em uma prole e, então, se separaram novamente, quando a prole produziu gametas, ao invés de se misturarem como um líquido contínuo. Mas Mendel não poderia ter deduzido esse mecanismo se tivesse estudado a variação da altura na maioria das espécies de planta, onde tal variação é contínua, pois ela depende de muitas diferenças gênicas. Assim, uma sorte na escolha do organismo-modelo pode possibilitar descobrir mecanismos de grande generalidade, mas uma escolha infeliz pode esconder tais mecanismos ou revelar fenômenos que existem apenas em alguns organismos.

A necessidade de uma variedade de organismos-modelo

Embora o uso de determinado organismo-modelo possa revelar características gerais da herança e desenvolvimento, não podemos saber o quanto essas características são gerais a menos que sejam feitos experimentos em uma variedade de características herdadas em uma variedade de organismos-modelo com padrões muito diferentes de reprodução e desenvolvimento.

Os organismos-modelo foram escolhidos, em parte, por suas propriedades básicas biológicas diferentes e, em parte, pelo pequeno tamanho dos indivíduos, curto tempo de geração e facilidade com que podem ser cultivados e cruzados sob condições simples controladas. Para o estudo da genética de vertebrados, os camundongos são preferidos aos elefantes.

A necessidade de estudar uma ampla gama de características biológicas e genéticas levou a uma gama de organismos-modelo para cada um dos grupos biológicos básicos (Figura 1.13).

Vírus Essas partículas simples não-vivas não têm toda a maquinaria metabólica. Um vírus infecta uma célula hospedeira e modifica seu aparelho biossintético para a produção de mais vírus, incluindo a replicação dos genes virais. Os vírus que infectam bactérias, chamados de *bacteriófagos*, são o modelo padrão (Figura 1.13a). O principal uso dos vírus foi estudar a estrutura física e química do DNA e os mecanismos fundamentais de replicação do DNA e mutação.

Procariontes Esses organismos vivos unicelulares não têm membrana nuclear nem compartimentos intracelulares. Embora exista uma forma especial de cruzamento e troca genética entre células procarióticas, na maior parte de suas vidas elas são essencialmente **haplóides** ao invés de **diplóides**. As células de organismos diplóides contêm dois conjuntos completos de genes, cada um tendo contribuído por um de seus dois genitores através de gametas que se fundem na fertilização. Os organismos haplóides têm apenas um conjunto de genes em cada célula.

A bactéria intestinal *Escherichia coli* é o modelo comum. Alguns geneticistas de *E. coli* estavam tão convictos da aplicabilidade geral de seu organismo-modelo que o correio do departamento da universidade imprimiu uma *E. coli* estilizada reestruturada para parecer um elefante.

Eucariontes Todas as outras formas celulares de vida são feitas de uma ou mais células com uma membrana nuclear e compartimentos celulares.

Leveduras Esses fungos unicelulares geralmente se reproduzem por divisão de células haplóides para formar colônias, mas também se reproduzem sexualmente pela fusão de duas células. O produto diplóide dessa fusão pode se reproduzir por divisão celular e formação de colônias, o que é depois seguido de meiose e produção de esporos haplóides que dão origem a novas colônias haplóides. *Saccharomyces cerevisiae* é a espécie-modelo usual.

Fungos filamentosos Nesses fungos, a divisão nuclear e o crescimento produzem longos filamentos ramificados chamados hifas. Elas são separadas irregularmente em “células” por membranas e paredes celulares, mas um desses compartimentos celulares pode conter mais de um núcleo haplóide. Uma fusão de dois filamentos resultará em um núcleo diplóide, que então sofre meiose para produzir um corpo frutífero de células haplóides. Nos fungos, a *Neurospora* é o organismo-modelo padrão (Figura 1.13b) porque seu corpo frutífero (veja Capítulo 3) contém oito esporos de forma linear, refletindo o pareamento de cromossomos e a síntese de novos filamentos cromossômicos na meiose.

A importância das bactérias, leveduras e fungos filamentosos para a Genética está em sua bioquímica básica. Para seu metabolismo e crescimento, eles precisam apenas de uma fonte de carbono, tal como o açúcar, de alguns minerais, tais como o cálcio, e, em alguns casos, de uma vitamina, tal como a biotina. Todos os outros componentes químicos da célula, incluindo todos os aminoácidos e nucleotídeos, são sintetizados por sua maquinaria celular. Assim, é possível estudar os efeitos de mudanças genéticas na maioria das vias bioquímicas básicas.

Organismos multicelulares Para o estudo genético da diferenciação das células, tecidos e órgãos, bem como desenvolvimento da forma do corpo, são necessários organismos mais complexos. Esses organismos devem ser fáceis de cultivar sob condições controladas, ter ciclos de vida curtos o suficiente para permitir experimentos de cruzamentos por muitas gerações e ser pequenos o suficiente para produzir um grande número de indivíduos práticos. Os principais organismos-modelo que preenchem esses requisitos são:

- *Arabidopsis thaliana*, uma pequena planta com flores que pode ser cultivada em grande número na estufa ou laboratório (Figura 1.13c). Ela tem um genoma pequeno contido apenas em cinco cromossomos. É um modelo ideal para estudar o desenvolvimento de plantas superiores e a comparação do desenvolvimento de animais e plantas e a estrutura do genoma.
- *Drosophila melanogaster*, uma mosca-das-frutas que tem apenas quatro cromossomos. No estágio larval, esses cromossomos têm um padrão bem marcante de bandas que possibilitam observar as mudanças físicas, tais com dele-

Um pequeno número de organismos-modelo é o foco da pesquisa genética



FIG. 1.13 Alguns organismos-modelo. (a) Bacteriófago λ ligado a uma *Escherichia coli* infectada; partículas de prole do fago estão amadurecendo dentro da célula. (b) *Neurospora* crescendo em uma árvore queimada após um incêndio na floresta. (c) *Arabidopsis*. (d) *Caenorhabditis elegans*. [(a) Lee D. Simon/Science Source/Photo Researchers; (b) cortesia de David Jacobson; (c) Wally Eberhart/Visuals Unlimited; (d) Sinclair Stammers/Photo Researchers.]

ções e duplicações, que podem ser correlacionadas com mudanças genéticas na morfologia e bioquímica. O desenvolvimento da *Drosophila* produz segmentos corpóreos em uma ordem ântero-posterior que exemplifica o plano corpóreo básico comum a invertebrados e vertebrados.

- *Caenorhabditis elegans*, um pequeno nematelminto com um total apenas de alguns milhares de células adultas. Essas células formam um sistema nervoso; um trato digestivo com uma boca, faringe e ânus; e um sistema reprodutivo que pode produzir tanto ovócitos quanto espermatozoides (Figura 1.13d).

- *Mus musculus*, o rato caseiro, o organismo-modelo para os vertebrados. O rato tem sido estudado para comparar a base genética do desenvolvimento de vertebrados e invertebrados, bem como para explorar a genética dos sistemas antígeno-anticorpo, de interações materno-fetais no útero e o envolvimento de genes no câncer.

Os genomas de todos os organismos-modelo precedentes foram seqüenciados. A despeito das grandes diferenças em biologia, existem muitas similaridades em seus genomas. A Figura 1.14 compara os genomas de eucariontes, procariontes e vírus.

Comparação estrutural dos componentes do genoma de eucariontes, procariontes e vírus

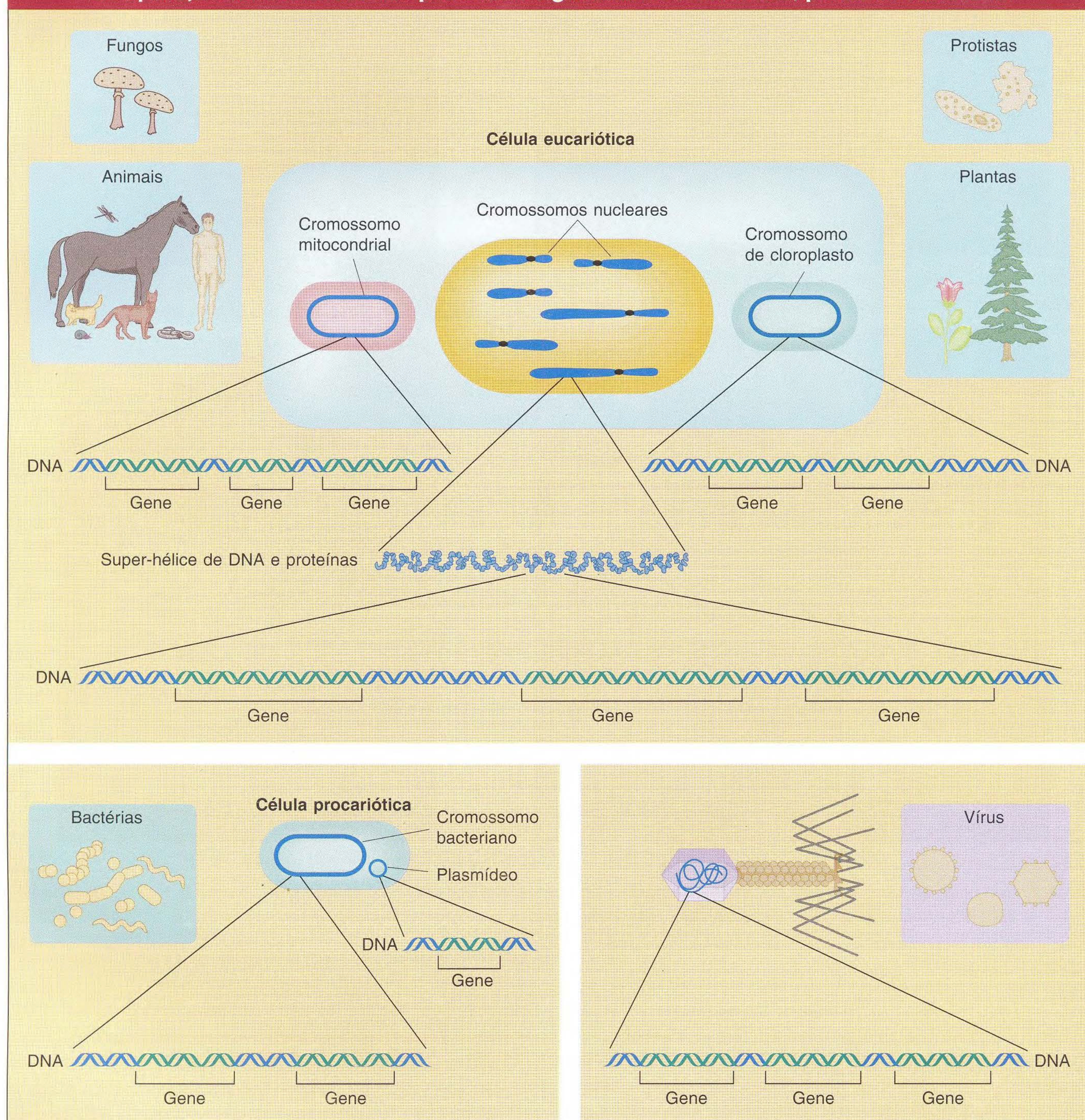


FIG. 1.14 Eucariontes, procariontes e vírus, todos contêm cromossomos nos quais residem genes, mas existem algumas diferenças nos genomas. Por exemplo, cromossomos procarióticos são circulares, enquanto os cromossomos virais e os nucleares eucarióticos são lineares. Duas organelas eucarióticas – as mitocôndrias e os cloroplastos – contêm cromossomos circulares separados.

A maioria dos estudos de genética são feitos em limitado número de organismos-modelo com características que os tornam especialmente adequados ao estudo científico.

► 1.6 Genes, o Ambiente e o Organismo

Os genes não podem ditar a estrutura de um organismo por si só. O outro componente crucial na fórmula é o ambiente. O ambiente influencia a ação gênica de muitos modos, sobre os quais aprenderemos nos capítulos subseqüentes. Mais concretamente, o ambiente fornece as matérias-primas para os processos sintéticos controlados pelos genes. Um broto torna-se um carvalho usando, no processo, apenas água, oxigênio, dióxido de carbono, alguns materiais inorgânicos do solo e energia da luz.

Modelo I: Determinação genética

Praticamente todas as diferenças entre as espécies são determinadas por diferenças em seus genomas. Não há ambiente no qual um leão dará origem a um cordeiro. Um broto desenvolve-se em um carvalho, enquanto um esporo de musgo desenvolve-se em um musgo, mesmo que ambos cresçam lado a lado na mesma floresta. As duas plantas que resultam desses processos desenvolvimentais assemelham-se a seus genitores e diferem uma da outra, muito embora tenham acesso à mesma restrita gama de materiais do meio.

Mesmo dentro das espécies, alguma variação é totalmente uma consequência de diferenças genéticas que não podem ser modificadas por qualquer mudança no que pensamos ser ambiental. Os filhos de escravos africanos trazidos para os EUA têm pele escura, inalterada pela mudança de seus genitores da África tropical para a temperada Maryland. A possibilidade de muitas experiências genéticas depende do fato de que muitas das diferenças fenotípicas entre indivíduos mutantes e tipo selvagem resultantes de diferenças alélicas são insensíveis a condições ambientais. O poder determinante dos genes é geralmente demonstrado por diferenças nas quais um alelo é normal e o outro anormal. A doença humana herdada anemia falciforme é um bom exemplo. A causa subjacente da doença é uma variante da hemoglobina, a proteína transportadora de oxigênio encontrada nas hemácias. As pessoas normais têm um tipo de hemoglobina chamada *hemoglobina A*, para a qual a informação está codificada em um gene. Uma única mudança de nucleotídeo no DNA desse gene, levando a uma mudança em um único aminoácido no polipeptídeo, resulta na produção de uma hemoglobina ligeiramente alterada, chamada *hemoglobina S*. Nas pessoas que possuem apenas hemoglobina S, o efeito final dessa pequena mudança no DNA é uma doença grave, a anemia falciforme, geralmente fatal.

Tais observações, se generalizadas, levam a um modelo, mostrado na Figura 1.15, de como os genes e o ambiente influenciam o fenótipo. Desse modo, os genes atuam como um conjunto de instruções para transformar materiais

Um modelo de determinação que enfatiza o papel dos genes

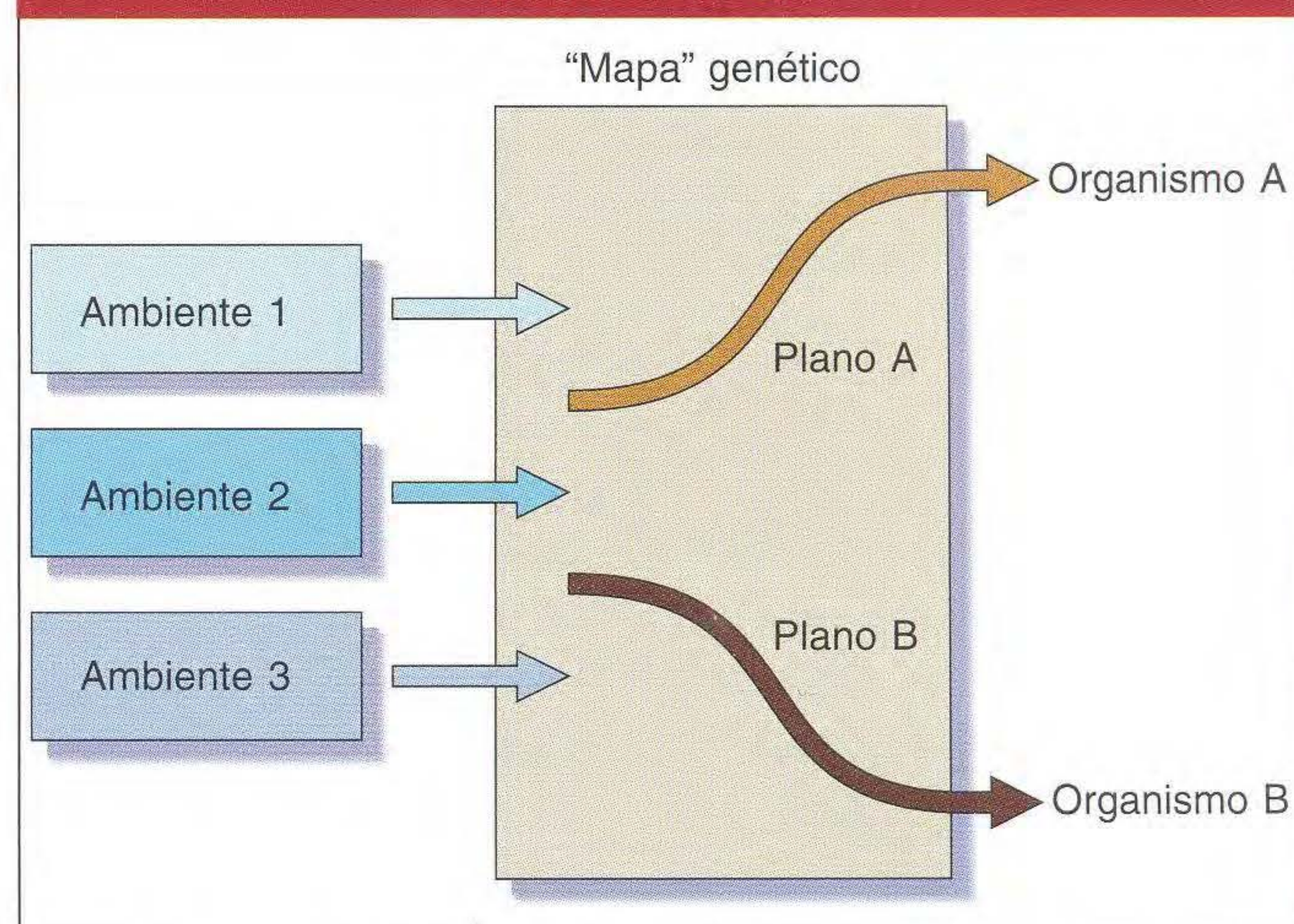


FIG. 1.15 Nesse modelo, os ambientes diferentes fornecem, todos, as condições ambientais básicas necessárias para o desenvolvimento, e as diferenças no fenótipo são uma consequência das diferenças na "constituição" genética.

ambientais mais ou menos indiferenciados em um organismo específico, assim como um mapa específica que forma de casa será construída a partir de materiais básicos. Os mesmos tijolos, madeiras etc. podem ser usados em uma casa em forma de A ou de teto achatado, de acordo com planos diferentes. Tal modelo significa que os genes são realmente os elementos dominantes na determinação fenotípica. O ambiente simplesmente fornece a matéria-prima indiferenciada.

Modelo II: Determinação ambiental

Considere duas gêmeas monozigóticas ("idênticas"), os produtos de um único ovócito fertilizado que se dividiu e produziu duas crianças completas com genes idênticos. Suponha que as gêmeas nasceram na Inglaterra, mas foram separadas ao nascimento e levadas para países diferentes. Se uma é criada na China por pais adotivos que falam chinês, ela irá falar chinês, enquanto a irmã criada em Budapeste falará húngaro. Cada uma absorverá os valores culturais de seu ambiente. Embora as gêmeas comecem a vida com propriedades genéticas idênticas, os ambientes culturais diferentes nos quais vivem produzirão diferenças entre elas (e diferenças de seus genitores). Obviamente, as diferenças nesse caso são devidas ao ambiente, e os efeitos genéticos não têm importância na determinação das diferenças.

Esse exemplo sugere o modelo da Figura 1.16, que é o contrário do mostrado na Figura 1.15. No modelo da Figura 1.16, os genes influenciam o sistema, dando certos sinais gerais para o desenvolvimento, mas o ambiente determina o curso real do desenvolvimento. Imagine um conjunto de especificações para uma casa que simplesmente peça um "chão que suporte 13.590 kg por pé ao quadrado" ou "paredes com um fator de isolamento de 15 polegadas"; o aspecto geral e outras características da estrutura seriam determinados pelos materiais de construção disponíveis.

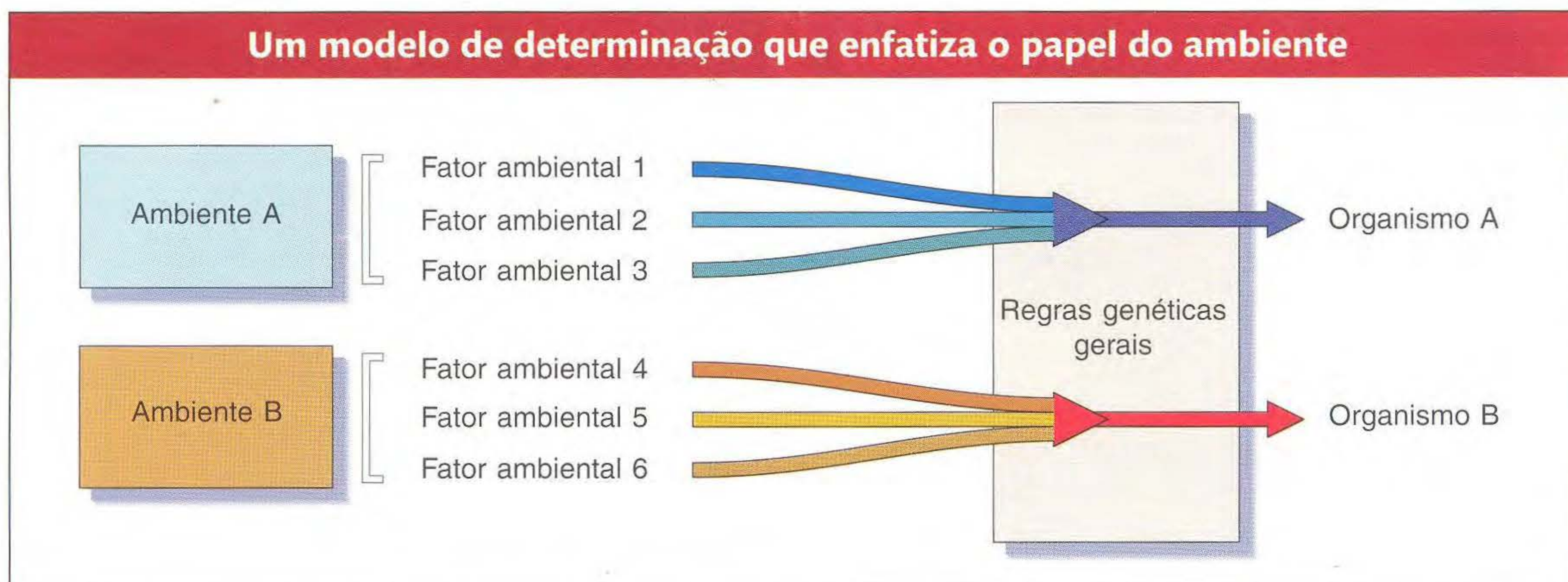


FIG. 1.16 Nesse modelo, os organismos A e B diferem principalmente porque se desenvolveram em ambientes diferentes.

Modelo III: Interação genótipo-ambiente

Em geral, lidamos com organismos que diferem tanto nos genes quanto no ambiente. Se desejamos prever como um organismo vivo irá desenvolver-se, precisamos saber tanto a constituição genética que ele herda de seus genitores (seu **genótipo**) quanto a sequência histórica de ambientes aos quais ele foi exposto. Cada organismo tem uma história de desenvolvimento desde a concepção até a morte. O que um organismo será no momento seguinte depende criticamente tanto de seu estado atual quanto do ambiente que ele encontra durante esse momento. Faz diferença para um organismo não só os ambientes que ele encontra, mas também em que sequência ele os encontra. Uma mosca-das-frutas (*Drosophila melanogaster*), por exemplo, desenvolve-se normalmente a 25°C. Se a temperatura for por pouco tempo elevada para 37°C em seu estágio de pupa do desenvolvimento, algumas das moscas adultas não terão parte do padrão de veia normal em suas asas. Entretanto, se esse “choque de temperatura” é administrado apenas 24 h depois, o padrão de veias desenvolve-se normalmente. Um modelo geral no qual os genes e o ambiente determinam em conjunto (por algumas regras de desenvolvimento) as características reais de um organismo é mostrado na Figura 1.17).

À medida que um organismo se transforma desenvolvimentalmente de um estágio para outro, seus genes interagem com seu ambiente em cada momento de sua história de vida. A interação de genes e ambiente determina o que os organismos são.

O uso do genótipo e fenótipo

À luz da discussão precedente, hoje podemos entender melhor o uso dos termos *genótipo* e *fenótipo*.

Um organismo típico assemelha-se a seus genitores mais do que se assemelha a membros de sua espécie aos quais não é aparentado. Assim, geralmente falamos como se as próprias características individuais fossem herdadas: “Ele herdou a inteligência da mãe” ou “Ela herdou diabetes do pai”. Mas a seção anterior mostra que tais frases são incorretas. “Sua inteligência” e “seu diabetes” desenvolvem-se por meio de longas sequências de eventos nas histórias de vida das pessoas afetadas, e tanto os genes quanto o ambiente têm papéis nessas sequências. No sentido biológico, um organismo herda apenas as estruturas moleculares do zigoto do qual se desenvolve. Organismos individuais herdam seus genes, não os produtos finais de suas histórias de desenvolvimento.

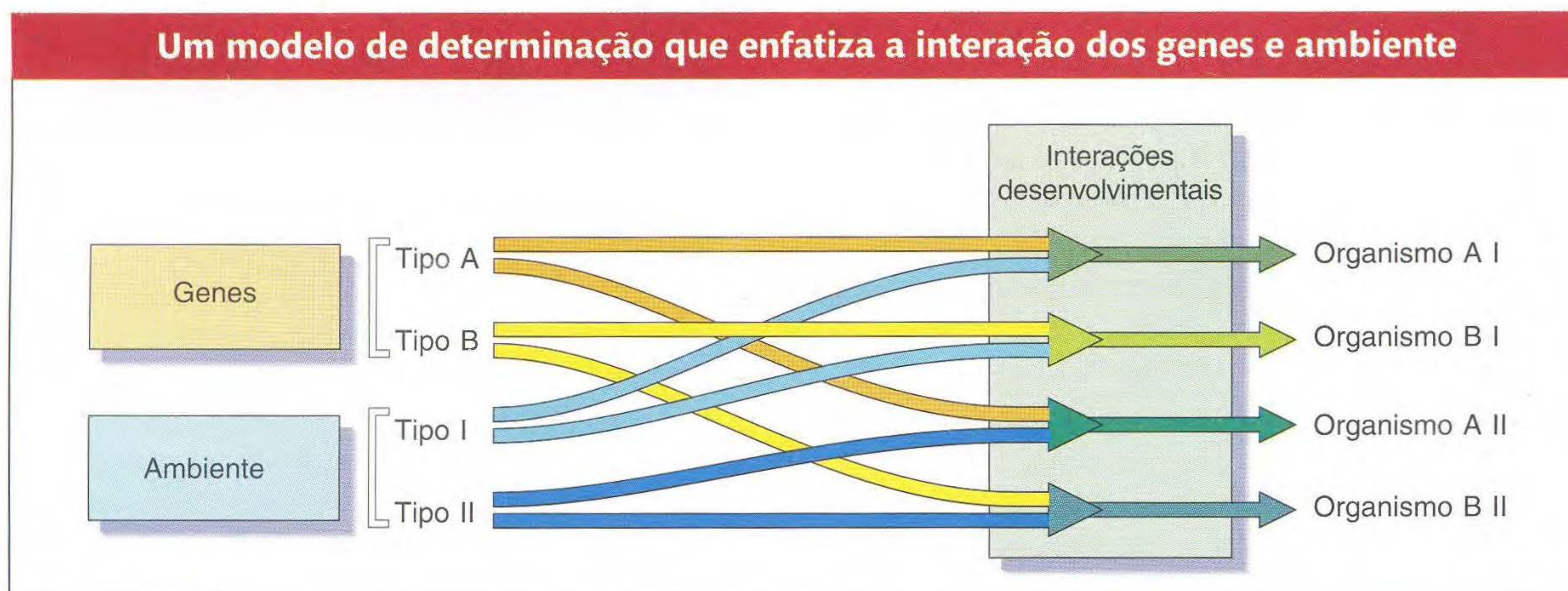


FIG. 1.17 Nesse modelo, os genes e o ambiente em conjunto determinam as características de um organismo.

Para evitar tal confusão entre genes (que são herdados) e os resultados desenvolvimentais (que não são), os geneticistas fazem a distinção fundamental entre o genótipo e o fenótipo de um organismo. Os organismos têm o mesmo genótipo em comum se tiverem o mesmo conjunto de genes. Os organismos têm o mesmo fenótipo se parecerem ou funcionarem iguais.

Estritamente falando, o genótipo descreve o conjunto completo de genes herdados por um organismo individual, e o fenótipo descreve todos os aspectos da morfologia, fisiologia, comportamento e relações ecológicas do organismo individual. Nesse sentido, não existem dois indivíduos que pertençam ao mesmo fenótipo, pois sempre existe alguma diferença (mesmo pequena) entre eles na morfologia ou fisiologia. Adicionalmente, exceto para indivíduos produzidos por outro organismo por reprodução assexual, quaisquer dois organismos diferem pelo menos um pouco no genótipo. Na prática, usamos os termos *genótipo* e *fenótipo* em um sentido mais restrito. Lidamos com alguma descrição fenotípica parcial (digamos, cor de olho) e com um subgrupo de genótipo (digamos, os genes que afetam a pigmentação do olho).

Quando usamos os termos *fenótipo* e *genótipo*, geralmente queremos dizer "fenótipo parcial" e "genótipo parcial", e especificamos uma ou algumas características e genes que são os subgrupos de interesse.

Note uma diferença muito importante entre genótipo e fenótipo: o genótipo é essencialmente uma característica fixa de um organismo individual; o genótipo permanece constante durante a vida, exceto por raras mutações nas células, e é essencialmente inalterado pelos efeitos ambientais. A maioria dos fenótipos muda continuamente durante a vida de um organismo à medida que seus genes interagem com uma sequência de ambientes. A fixação do genótipo não significa a fixação do fenótipo.

Ruído de desenvolvimento

Até agora, supomos que um fenótipo é unicamente determinado pela interação de um genótipo específico e um ambiente específico. Mas uma visão mais atenta mostra alguma variação a mais inexplicada. A mosca-das-frutas *Drosophila*, como todos os insetos, tem um olho composto feito de um grande número de receptores de luz chamados facetas oculares (Figura 1.18a). Existem mutações que reduzem o número dessas facetas, duas das quais, *infrabar* e *ultrabar*, são mostradas na Figura 1.18b em contraste com o tipo selvagem. O número de facetas oculares também é afetado pela temperatura na qual a mosca se desenvolve (Figura 1.18c). Uma *Drosophila* do genótipo tipo selvagem criada a 16°C tem 1.000 facetas em cada olho, enquanto tem apenas 800 facetas se desenvolvida a 25°C. De fato, esses valores são apenas médias; uma mosca criada a 16°C pode ter 980 facetas, e outra pode ter 1.020. Talvez essas variações sejam devidas a pequenas flutuações no ambiente local ou pequenas diferenças nos genótipos. Entretanto, uma contagem típica pode mostrar

que uma mosca tipo selvagem tem, digamos, 1.017 facetas no olho esquerdo e 982 no olho direito. Em outra mosca tipo selvagem criada nas mesmas condições experimentais, o olho esquerdo tem um pouco menos facetas que o olho direito, embora o olho esquerdo e direito da mesma mosca sejam geneticamente idênticos. Além disso, sob típicas condições experimentais, a mosca desenvolve-se como uma larva (alguns milímetros de tamanho), alimentando-se de alimentos artificiais homogêneos em uma garrafa de laboratório e completando seu desenvolvimento como uma pupa (também com alguns milímetros de tamanho) grudada verticalmente ao lado interno da garrafa acima da superfície do alimento. Certamente, o ambiente não difere significativamente de um lado da mosca para o outro. Se os dois olhos estão submetidos à mesma sequência de ambientes e são geneticamente idênticos, então por que existe diferença fenotípica entre os olhos esquerdo e direito?

As diferenças em forma e tamanho dependem parcialmente do processo de divisão celular que transforma o zigoto em um organismo multicelular. A divisão celular, por sua vez, é sensível a eventos moleculares dentro da célula, e esses eventos podem ter um componente aleatório relativamente grande. Por exemplo, a vitamina biotina é essencial para o crescimento da *Drosophila*, mas sua concentração média é apenas de uma molécula por célula. A taxa de qualquer processo que dependa da presença dessa molécula irá flutuar à medida que a concentração varia. Menos facetas oculares podem se desenvolver se a disponibilidade de biotina por acaso flutuar para menos dentro de um curto período desenvolvimental durante o qual o olho está sendo formado. Assim, esperaríamos uma variação aleatória em tais características fenotípicas, como o número de células oculares, o número de cabelos, a forma exata de características pequenas e as conexões de neurônios em um complexo sistema nervoso central quando o genótipo e o ambiente estão precisamente fixados. Os eventos aleatórios no desenvolvimento levam a uma variação no fenótipo chamada de **ruído desenvolvimental**.

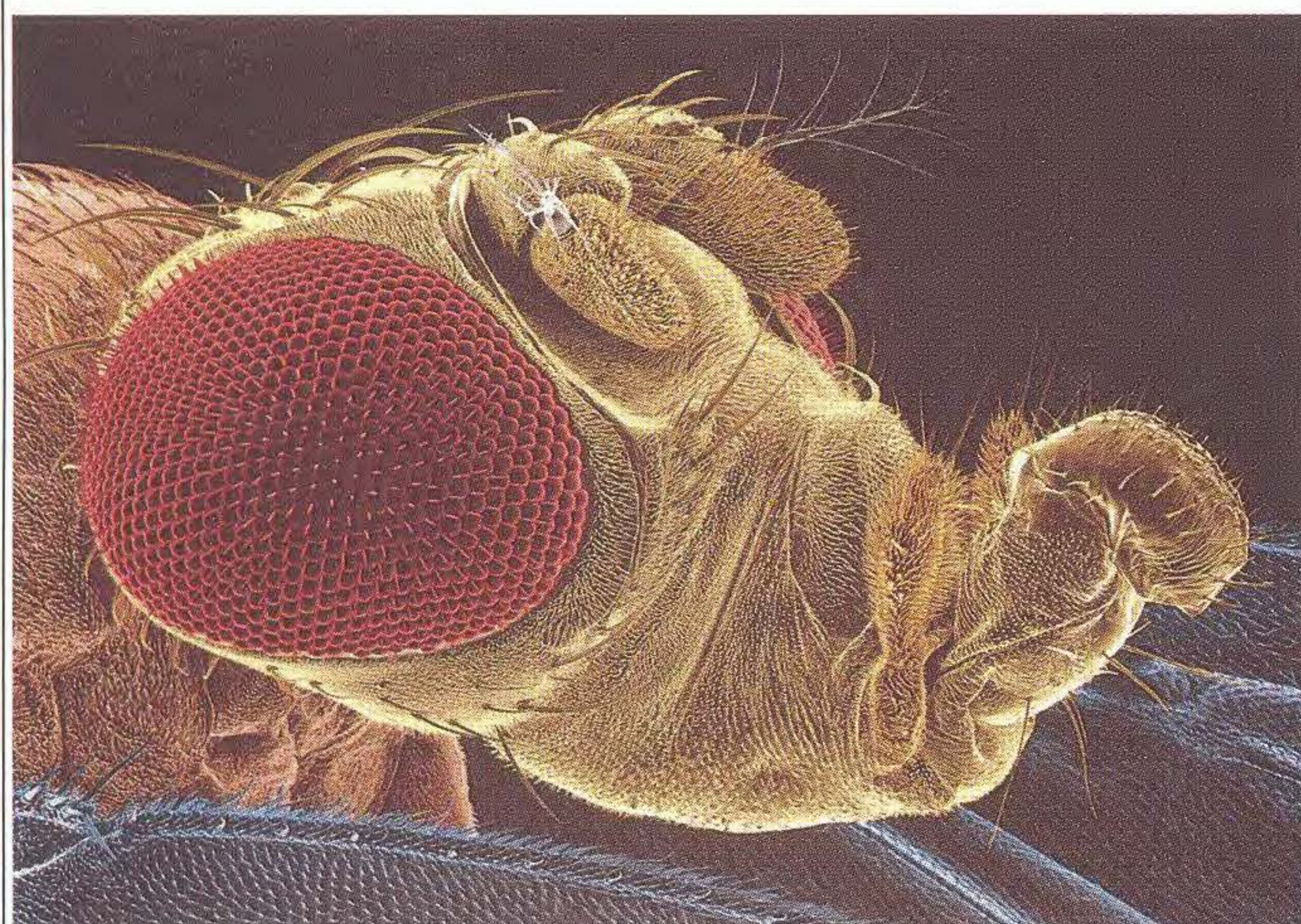
Em algumas características, tais como as células oculares em *Drosophila*, o ruído desenvolvimental é uma fonte importante das variações observadas no fenótipo.

Adicionando o ruído desenvolvimental a nosso modelo de desenvolvimento fenotípico, obtemos algo como a Figura 1.19. Com determinado genótipo e ambiente, há uma gama de possíveis resultados para cada etapa desenvolvimental. O processo desenvolvimental contém sistemas de *feedback* que tendem a manter os desvios dentro de alguns limites, de modo que a gama de desvios não aumenta indefinidamente nas muitas etapas do desenvolvimento. Entretanto, esse *feedback* não é perfeito. Para qualquer determinado genótipo que se desenvolva em determinada sequência de ambientes, resta alguma incerteza quanto ao fenótipo exato que resultará.

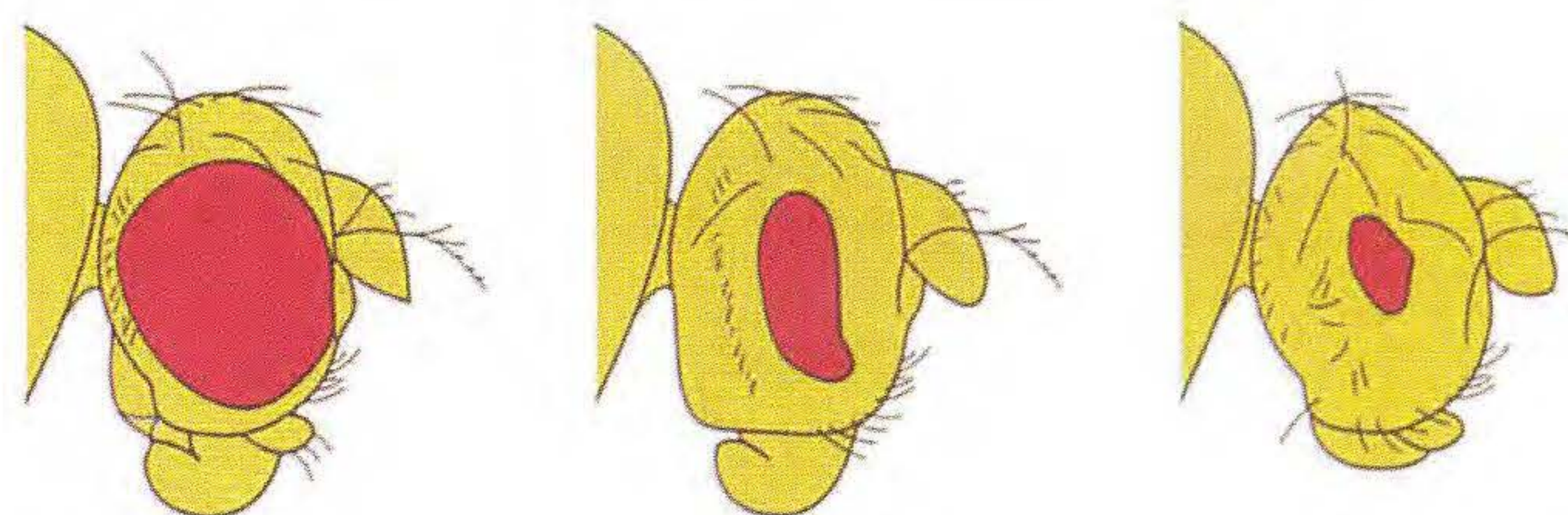
Três níveis de desenvolvimento

O Capítulo 12 está envolvido com o modo pelo qual os genes medeiam o desenvolvimento, mas em nenhuma parte desse

Genes e ambiente, ambos influenciam o tamanho do olho em *Drosophila*



(a)

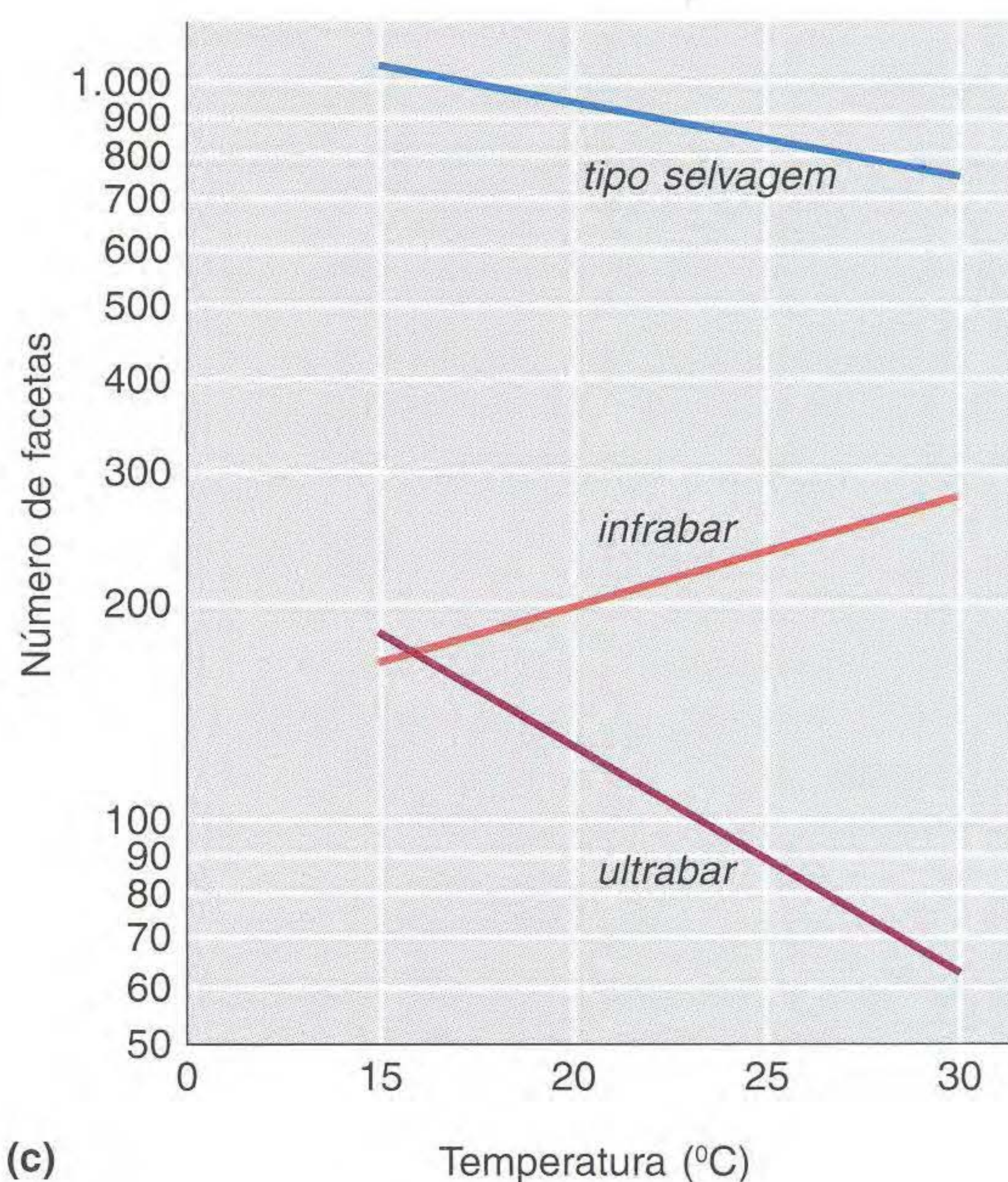


Tipo selvagem

Infrabar

Ultrabar

(b)



(c)

FIG. 1.18 Variação genética e ambiental do tamanho do olho de *Drosophila*. (a) Mostrando como o olho normal é feito de centenas de unidades chamadas facetas. O número de facetas determina o tamanho do olho. (b) Tamanhos relativos dos olhos dos tipos selvagem, infrabar e ultrabar. (c) Efeito da temperatura no tamanho do olho para os três genótipos. [(a) Don Rio e Sima Misra, University of California, Berkeley.]

Ruído desenvolvimental é outro fator que influencia o fenótipo

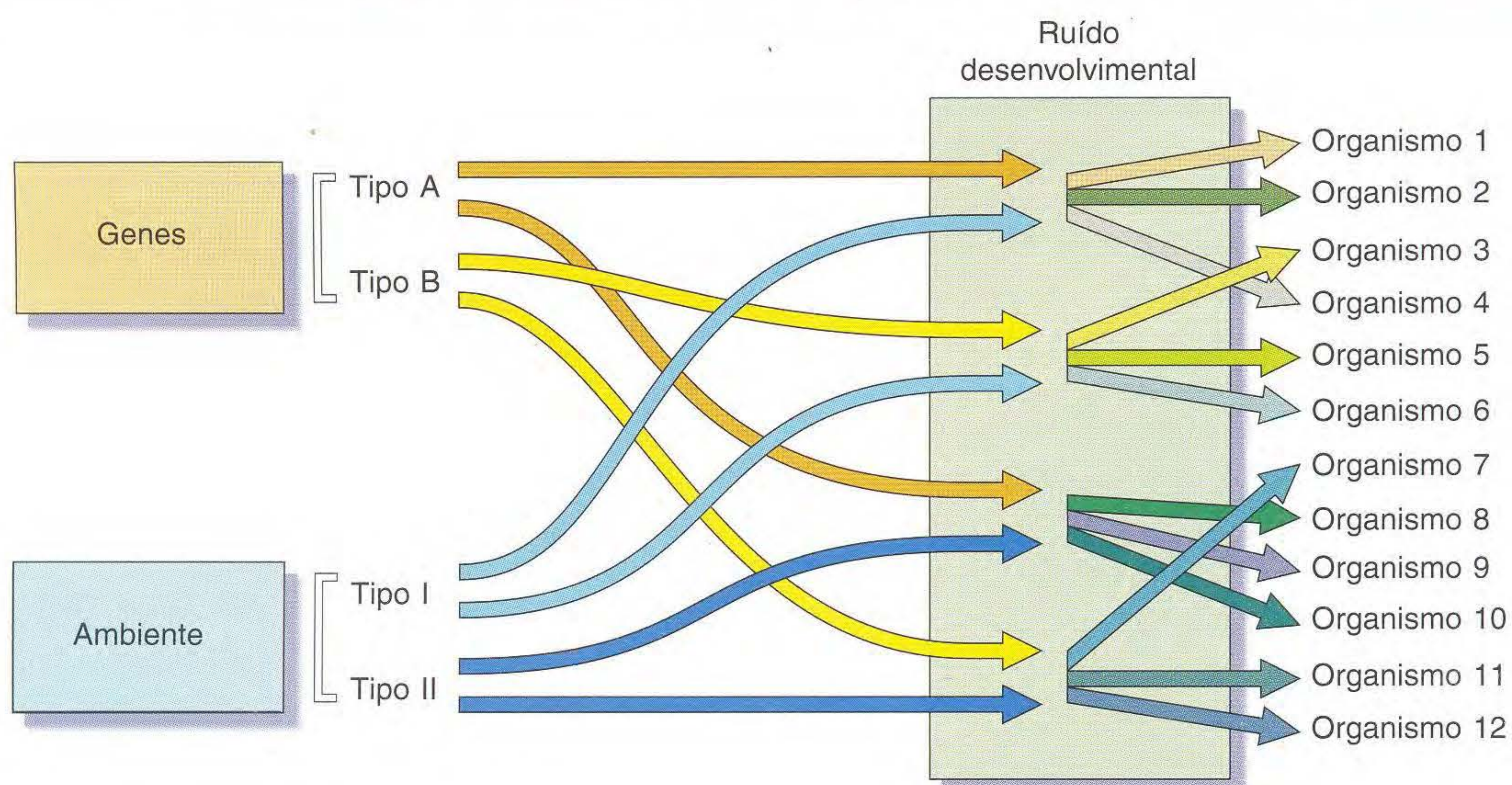


FIG. 1.19 Um modelo de determinação fenotípica que mostra como os genes, o ambiente e o ruído desenvolvimental interagem para produzir um fenótipo.

capítulo consideramos o papel do ambiente ou a influência do ruído desenvolvimental. Como pudemos, no começo deste livro, enfatizar o papel conjunto dos genes, ambiente e ruído em influenciar o fenótipo, embora, em nossa posterior consideração do desenvolvimento, ignorássemos o ambiente e o ruído? A resposta é que a genética desenvolvimental moderna está envolvida com processos muito básicos de diferenciação que são comuns a todos os membros individuais de uma espécie e, de fato, são comuns a animais tão diferentes quanto as moscas-das-frutas e os mamíferos. Como a parte frontal de um animal se torna diferenciada da parte posterior, o lado ventral do lado dorsal? Como o corpo se torna segmentado, e por que os membros se formam em alguns segmentos e não em outros? Por que os olhos se formam na cabeça e não na metade do abdome? Por que as antenas, asas e pernas de uma mosca parecem tão diferentes, muito embora sejam derivadas na evolução de apêndices que pareciam similares nos ancestrais mais antigos dos insetos? Nesse nível de desenvolvimento, que é constante através dos indivíduos e espécies, a variação ambiental não tem papel, e podemos falar corretamente de genes “determinando” o fenótipo. Exatamente porque os efeitos dos genes podem ser isolados nesse nível de desenvolvimento e exatamente porque os processos parecem ser gerais através de uma ampla variedade de organismos, eles são mais fáceis de estudar do que as características para as quais as variações ambientais são importantes, e a genética do desenvolvimento concentrou-se em compreendê-las.

Em um segundo nível de desenvolvimento, existem variações nos temas desenvolvimentais básicos que são diferentes entre espécies mas são constantes dentro das espécies, e essas variações, também, podem ser compreendidas concentrando-se nos genes, embora no momento elas não sejam parte do estudo da genética do desenvolvimento. Assim, embora tanto leões quanto ovelhas tenham quatro pernas, uma em cada canto, os leões sempre dão origem a leões e as ovelhas a ovelhas, e nós não temos dificuldades em distingui-los em qualquer ambiente em que estiverem. Novamente, podemos dizer que os genes “determinam” a diferença entre as duas espécies, embora tenhamos que ter mais cautela aqui. Duas espécies podem diferir em algumas características porque vivem em ambientes muito diferentes, e, até que possamos criá-las no mesmo ambiente, nem sempre podemos estar certos de que a influência ambiental tem um papel. Por exemplo, duas espécies de *baboons* na África, uma vivendo em locais muito secos da Etiópia e outra nas áreas mais produtivas de Uganda, têm um comportamento muito diferente de obter alimento e estrutura social. Sem transplantar colônias das duas espécies entre os dois ambientes, não podemos saber o quanto da diferença é uma resposta direta desses primatas a diferentes condições alimentares.

É nesse terceiro nível, as diferenças na morfologia, fisiologia e comportamento entre os indivíduos dentro de uma espécie, que os fatores genéticos, ambientais e ruído desenvolvimental são discutidos neste capítulo. Um dos erros mais sérios na compreensão da genética pelos não-geneticistas tem sido a confusão entre a variação nesse nível e a variação no primeiro e segundo níveis. Os experimentos e descobertas a serem discutidos no Capítulo 12 não são, e não pretendem ser, modelos para a causa da variação individual. Eles se aplicam diretamente apenas a essas características, deliberadamente escolhidas, que são características gerais do desenvolvimento, e para as quais a variação ambiental parece ser irrelevante.

Os indivíduos dentro de uma espécie, que os fatores genéticos, ambientais e ruído desenvolvimental são discutidos neste capítulo. Um dos erros mais sérios na compreensão da genética pelos não-geneticistas tem sido a confusão entre a variação nesse nível e a variação no primeiro e segundo níveis. Os experimentos e descobertas a serem discutidos no Capítulo 12 não são, e não pretendem ser, modelos para a causa da variação individual. Eles se aplicam diretamente apenas a essas características, deliberadamente escolhidas, que são características gerais do desenvolvimento, e para as quais a variação ambiental parece ser irrelevante.

► Resumo

A Genética é o estudo dos genes em todos os níveis, desde as moléculas até as populações. Como uma disciplina moderna, ela começou nos anos de 1860 com o trabalho de Gregor Mendel, que primeiro formulou a idéia da existência de genes. Hoje, sabemos que um gene é uma região funcional da longa molécula de DNA que constitui a estrutura fundamental de um cromossomo. O DNA é composto de quatro nucleotídeos, cada um contendo um açúcar desoxirribose, fosfato e uma de quatro bases: adenina (A), timina (T), guanina (G) e citosina (C). O DNA é uma cadeia de nucleotídeos mantidos juntos por pontes entre A e T e entre G e C. Na replicação, as duas cadeias de nucleotídeos separam-se, e suas bases expostas são usadas como moldes para a síntese de duas moléculas-filhas idênticas de DNA.

A maioria dos genes codificam a estrutura de uma proteína (as proteínas são os principais determinantes das propriedades de um organismo). Para fazer uma proteína, o DNA primeiro é transcrito pela enzima RNA polimerase em uma cópia unifilamentar chamada RNA mensageiro. A sequência de nucleotídeos no mRNA é traduzida em uma sequência de aminoácidos que constitui a estrutura primária de uma proteína. As cadeias de aminoácidos são sintetizadas

nos ribossomos. Cada aminoácido é levado para o ribossomo por uma molécula de RNA transportador que se liga por sua trinca (o anticódon) a uma trinca códon no mRNA.

O mesmo gene pode ter formas alternativas, ou variantes. Os indivíduos podem ser classificados pelo genótipo (as variantes gênicas presentes) ou pelo fenótipo (características observáveis de aspecto ou fisiologia). Tanto os genótipos quanto os fenótipos apresentam variação dentro de uma população.

A análise dessas variantes fornece um poderoso método de investigar as propriedades biológicas em geral. Quando as variantes da propriedade de interesse não existem naturalmente, elas podem ser criadas introduzindo mudanças na informação genética, chamadas mutações. Existem dois enfoques para a análise genética: direto e reverso. A análise genética direta começa com a variação observada em uma propriedade morfológica ou fisiológica. Fazendo cruzamentos e procurando padrões de herança na prole, um pesquisador pode identificar genes que contribuem para a propriedade. A pesquisa continua para determinar como a informação contida no gene é normalmente transformada em eventos celulares e fisiológicos. A análise genética reversa depende da recente

habilidade em determinar a sequência de DNA de um genoma inteiro. O pesquisador começa alterando a sequência de DNA de um modo específico e, então, procura os resultados.

No laboratório, os geneticistas estudam a variação fenotípica sob condições ambientais controladas nas quais há uma correspondência de um a um entre as diferenças fenotípicas

e as diferenças genéticas. Entretanto, quando há variação no ambiente no qual os organismos se desenvolvem e funcionam, como na natureza, a relação entre genótipo e fenótipo é mais complexa. Para cada genótipo, há uma gama de fenótipos que podem surgir, devido à variação ambiental e porque o desenvolvimento está sujeito a variação molecular aleatória.

► Termos-chave

ácido desoxirribonucleico (DNA)	genética direta	RNA mensageiro (mRNA)
ácido ribonucleico (RNA)	genética reversa	RNA transportador (tRNA)
adenina (A)	genoma	ruído desenvolvimental
anticódon	genômica	sonda
base nucleotídica	genômica comparativa	timina (T)
bases complementares	genótipo	tipo selvagem
bioinformática	guanina (G)	tradução
citosina (C)	haplóide	transcrição
códon	mutação	transcrito
diplóide	mutante	transferência de Southern
enzima de restrição	organismo-modelo	transferência Northern
eucarionte	polimorfismo genético	transferência Western
fenótipo	procarionte	vírus
gene	proteína	
genética	ribossomo	

► Problemas

PROBLEMAS BÁSICOS

1. Defina *genética*. Você acha que os antigos egípcios criadores de cavalos eram geneticistas? Como suas abordagens diferiam das dos geneticistas modernos?
2. Como o DNA dita as propriedades gerais de uma espécie?
3. Quais são as características do DNA que lhe permitem fazer seu papel como molécula hereditária? Você pode pensar em tipos alternativos de moléculas hereditárias que podem ser encontradas em formas de vida extraterrestres?
4. Quantas moléculas diferentes de DNA com 10 pares de nucleotídeos são possíveis?
5. Se a timina constitui 15% das bases em certa amostra de DNA, que percentagem de bases deverão ser citosinas?
6. O conteúdo G + C de uma amostra de DNA é de 48%. Quais as proporções dos quatro nucleotídeos diferentes?
7. Cada célula do corpo humano contém 46 cromossomos.
 - a. Quantas moléculas de DNA essa afirmativa representa?
 - b. Quantos tipos diferentes de moléculas de DNA ela representa?
8. Um certo segmento de DNA tem a seguinte sequência de nucleotídeos em um filamento:

ATTGGTGCATTACTTCAGGCTCT

Qual deve ser a sequência do outro filamento?

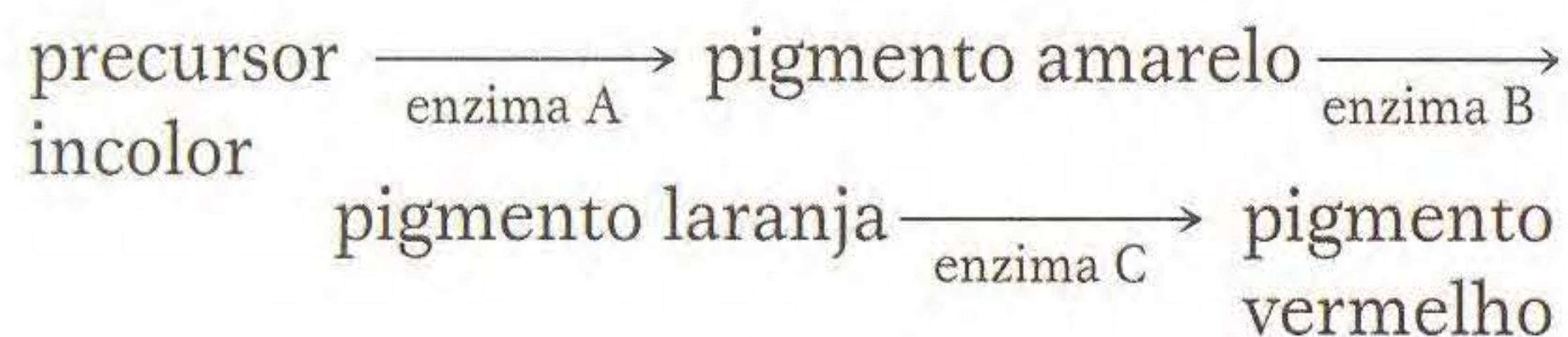
9. Em um único segmento de DNA, é possível que o número de adeninas seja maior que o número de timinas?

10. Em uma dupla hélice normal de DNA, é verdade que
 - a. A mais C seja sempre igual a G mais T?
 - b. A mais G seja sempre igual a C mais T?
11. Suponha que a seguinte molécula de DNA se replique para produzir duas moléculas-filhas. Escreva a sequência de nucleotídeos dessas moléculas-filhas usando preto para os nucleotídeos recém-polimerizados.

TTGGCACGTCGTAAT
AACCGTGCAGCATTA

12. Na molécula de DNA no Problema 11, suponha que o filamento de baixo é o filamento-molde e escreva o RNA transcrito.
13. Desenhe transferências Northern e Western de três variantes genéticas na Figura 1.8. (Suponha que a sonda usada na transferência Northern seja um clone do gene de tirosinase.)
14. O que é um gene? Quais são alguns dos problemas com sua definição?
15. O DNA é extraído de células de *Neurospora*, um fungo que tem um conjunto de sete cromossomos; ervilha, uma planta que tem dois conjuntos de sete cromossomos; e borboleta, um animal que tem dois conjuntos de seis cromossomos. Se for usada uma poderosa eletroforese para separar o DNA em um gel, quantas bandas cada uma dessas três espécies produzirá?
16. Se um códon no mRNA é UUA, qual o anticódon do tRNA que irá ligar-se a esse códon?

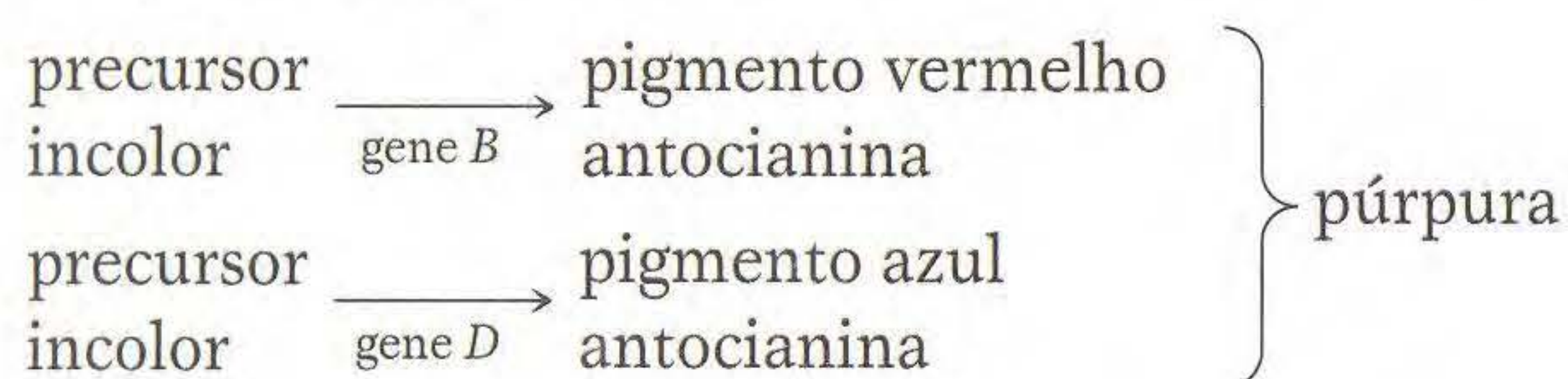
17. Duas mutações surgem em culturas separadas de um fungo normalmente vermelho (que tem apenas um conjunto de cromossomos). As mutações são encontradas em diferentes genes. A mutação 1 dá uma cor laranja, e a mutação 2 dá uma cor amarela. Os bioquímicos que trabalham na síntese do pigmento vermelho nessa espécie já descreveram a seguinte via:



- Que enzima está defeituosa no mutante 1?
- Que enzima está defeituosa no mutante 2?
- Qual seria a cor de um mutante duplo (1 e 2)?

PROBLEMAS DESAFIADORES

18. Em ervilhas, a cor púrpura das pétalas é controlada por dois genes, *B* e *D*. A via é



- Que cor de pétalas seria esperada em uma planta que leva duas cópias de uma mutação nula para *B*? (Uma mutação nula é uma mutação que resulta em um gene variante que falha em codificar uma proteína funcional.)
- Que cor de pétalas você esperaria em uma planta que leva duas cópias de uma mutação nula para *D*?
- Que cor de pétalas você esperaria em uma planta que é mutante duplo; isto é, leva duas cópias de uma mutação nula tanto para *B* quanto *D*?



Solucionando o Problema 18

- O que são ervilhas-de-cheiro, e como elas diferem das ervilhas comestíveis?
- O que é uma via no sentido usado aqui?
- Quantas vias são evidentes nesse sistema?
- As vias são independentes?
- Defina o termo *pigmento*.
- O que *incolor* significa nesse problema? Pense em um exemplo de algum soluto que seja incolor.
- Qual seria o aspecto de uma pétala contendo apenas substâncias incolores?
- A cor nas ervilhas-de-cheiro é algo parecido com uma mistura de tintas?
- O que é uma mutação?
- O que é uma mutação nula?
- Qual pode ser a causa de uma mutação nula ao nível de DNA?
- O que significa a expressão “duas cópias”? (Quantas cópias dos genes as ervilhas-de-cheiro normalmente têm?)
- Qual a relevância das proteínas para esse problema?
- Importa se os genes *B* e *D* estão no mesmo cromossomo?

- Escreva a representação do gene variante tipo selvagem de *B* e um mutante nulo a nível de DNA.
- Repita a questão 15 para o gene *D*.
- Repita a questão 15 para o mutante duplo.
- Como você explica a determinação genética da cor da pétala em ervilhas-de-cheiro para um jardineiro sem treinamento científico?
- É obtido um camundongo albino mutante cujo pigmento não tem melanina, normalmente feita por uma enzima *T*. O tecido do mutante não tem toda a atividade detectável para a enzima *T*. Entretanto, a transferência Western mostra claramente que uma proteína com propriedades imunológicas idênticas às da enzima *T* está presente nas células do mutante. Como isso é possível?
- Na Noruega, em 1934, uma mãe com dois filhos com retardo mental consultou o médico Asbjørn Følling. Durante a entrevista, Følling percebeu que a urina das crianças tinha um odor estranho. Mais tarde, ele testou a urina com cloreto férrico e descobriu que, enquanto a urina normal dá uma cor marrom, a urina das crianças corava em verde. Ele deduziu que a substância química responsável deveria ser o ácido fenilpirúvico. Devido à sua similaridade com a fenilalanina, parecia provável que essa substância tinha sido formada da fenilalanina no sangue, mas não havia teste para fenilalanina. Entretanto, determinada bactéria podia converter fenilalanina em ácido fenilpirúvico; assim, o nível de fenilalanina podia ser dosado usando o teste de cloreto férrico. As crianças foram encontradas tendo níveis altos de fenilalanina no sangue, e a fenilalanina era provavelmente a fonte do ácido fenilpirúvico. Essa doença, que veio a ser conhecida como fenilcetonúria (PKU), foi demonstrada como sendo herdada e causada por um alelo recessivo.

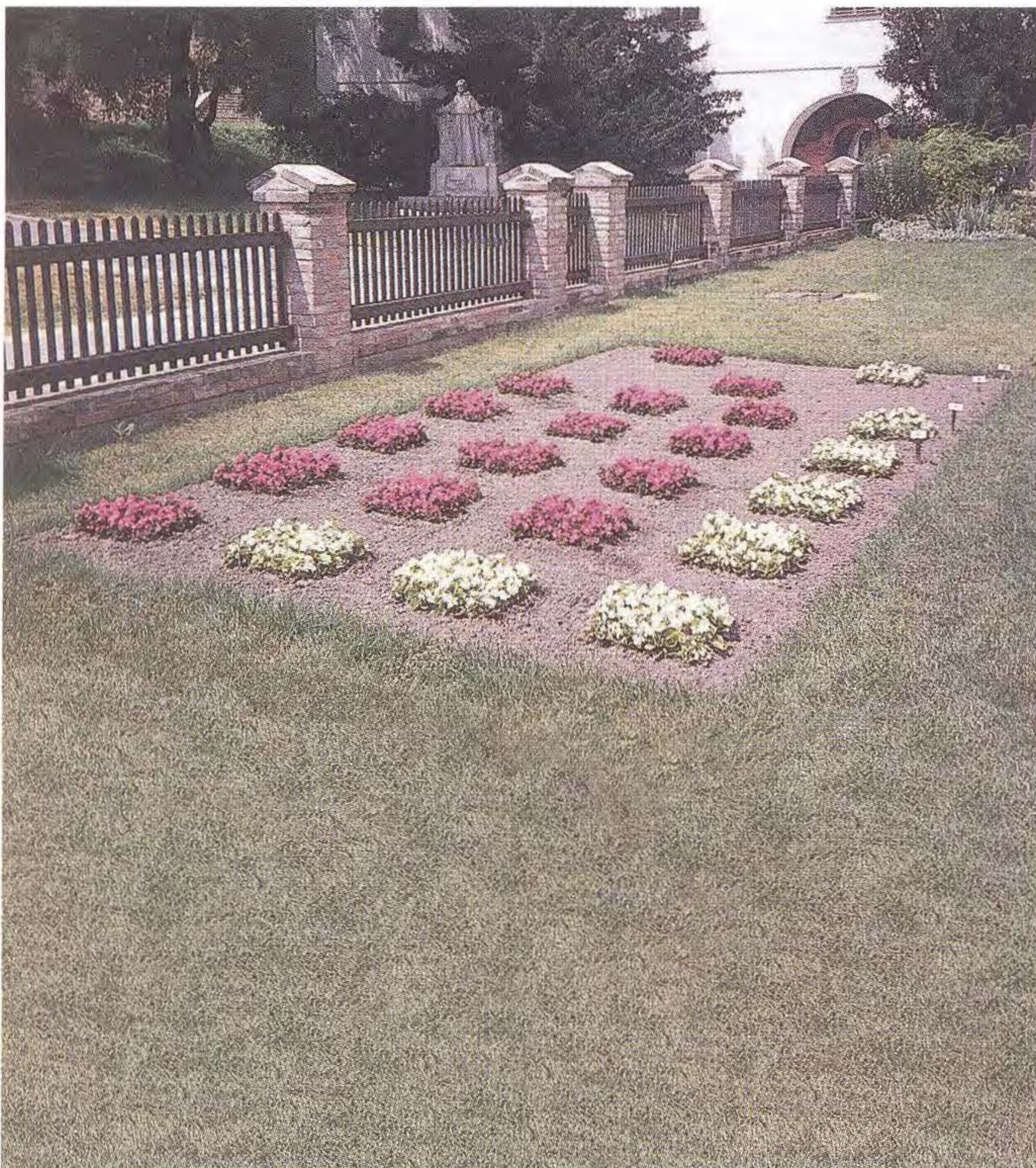
Ficou claro que a fenilalanina era a responsável, e que essa substância acumulava-se nos pacientes PKU e tinha sido convertida em altos níveis de ácido fenilpirúvico, que então interferiu no desenvolvimento normal do tecido nervoso. Esse achado levou à formulação de uma dieta especial pobre em fenilalanina, que podia alimentar os neonatos diagnosticados com PKU e que permitiam o desenvolvimento normal continuado sem o retardo. De fato, após o sistema nervoso da criança ter se desenvolvido, o paciente podia deixar a dieta especial. Entretanto, tragicamente, muitas mulheres PKU que se desenvolveram normalmente com a dieta especial tiveram filhos que nasceram com retardo mental, e a dieta especial não teve efeito nessas crianças.

- Por que você acha que os filhos de mães PKU nasceram retardados?
- Por que a dieta especial não teve efeito neles?
- Explique o motivo para a diferença nos resultados entre as crianças PKU e os filhos de mães PKU.
- Proponha um tratamento que permita que as mães PKU tenham filhos não afetados.

- e. Escreva um artigo curto sobre a PKU, integrando os conceitos nos níveis genético, diagnóstico, enzimático, fisiológico, de heredograma e populacional.
21. Compare e contraste os processos pelos quais a informação toma forma em um organismo e na construção de uma casa.
22. Tente pensar em exceções à seguinte frase: “Quando você observa um organismo, o que você enxerga é uma proteína ou algo que foi feito por uma proteína.”
23. A fórmula “genótipo + ambiente = fenótipo” é correta?

Herança Monogênica

2



O monastério do pai da Genética, Gregor Mendel. Uma estátua de Mendel é visível ao fundo. Hoje em dia, essa parte do monastério é um museu, e os curadores plantaram begônias vermelhas e brancas em uma disposição que representa graficamente o tipo de padrões de herança obtido por Mendel com as ervilhas. [Anthony Griffiths.]

Que tipos de pesquisas os biólogos fazem? A área central de pesquisa na biologia de todos os organismos envolve o desenvolvimento, tentando compreender o programa pelo qual um organismo desenvolve-se de um zigoto em um adulto, em outras palavras, o que torna um organismo do jeito que é. Geralmente, essa meta geral é subdividida no estudo das propriedades biológicas individuais, tais como o desenvolvimento da cor da flor da planta, ou locomoção animal, ou captação de um nutriente. Como os geneticistas analisam o desenvolvimento? Aprendemos, no Capítulo 1, que o enfoque genético para a compreensão de qualquer propriedade biológica é encontrar o subgrupo de genes no genoma que influencia essa propriedade, um processo às vezes chamado de **descoberta do gene**. Após esses genes terem sido identificados, o modo pelo qual os genes atuam para determinar a propriedade biológica pode ser elucidado por meio de mais pesquisas.

Perguntas Fundamentais

- Como são identificados os genes individuais?
- Que princípio de herança gênica Mendel descobriu?
- Qual a base cromossômica do princípio de Mendel?
- Como o princípio de Mendel é aplicado à herança humana?

Tópicos

- 2.1 Genes e cromossomos
- 2.2 Padrões de herança monogênica
- 2.3 A base cromossômica dos padrões de herança monogênica
- 2.4 Descobrendo genes pela observação das proporções de segregação
- 2.5 Padrões de herança monogênica ligada ao sexo
- 2.6 Análise de heredogramas

Existem vários tipos diferentes de enfoques analíticos para a descoberta do gene, mas um dos mais amplamente usados é baseado na detecção de *padrões de herança monogênica*, o tópico deste capítulo. Tais padrões de herança podem ser reconhecidos na prole de alguns tipos de reproduções controladas, que os geneticistas chamam de **cruzamentos**. Os componentes centrais desse tipo de análise são os **mutantes**, organismos individuais com alguma forma alterada de uma propriedade normal. A forma normal de qualquer propriedade de um organismo é chamada de **tipo selvagem**, que é encontrada “na natureza”. O *modus operandi* é cruzar um indivíduo apresentando a propriedade em sua forma tipo selvagem (por exemplo, uma planta com flores vermelhas) com um indivíduo apresentando uma forma mutada da propriedade (por exemplo, uma planta com flores brancas). A prole desse cruzamento é entrecruzada e, em sua prole, a proporção de plantas com flores vermelhas em relação àsquelas com flores brancas revelará se um único gene controla essa diferença na prole em estudo — nesse exemplo, vermelho *versus* branco. Por dedução, o tipo selvagem seria codificado pela forma tipo selvagem do gene, e a forma mutante seria codificada por uma forma do mesmo gene na qual o evento de mutação alterou de alguma forma a sequência de DNA. Outros mutantes afetando a cor da flor (talvez malva, manchada, com faixas etc.) seriam analisados do mesmo modo, resultando em um conjunto definido de “genes de cor de flores”. O uso de mutantes desse modo às vezes é chamado de **dissecção genética**, porque a propriedade biológica em questão (nesse caso, cor da flor) é separada para revelar seu programa genético subjacente não com um bisturi, mas com mutantes. Cada mutante potencialmente identifica um gene separado afetando essa propriedade.

Assim, cada projeto de descoberta de gene começa com uma caçada por mutantes que afetam o processo biológico que é o foco da pesquisa. O modo mais direto para obter mutantes é *triar* visualmente um grande número de indivíduos, procurando uma ocorrência ao acaso de mutantes nessa popula-

ção. Como exemplos, alguns dos resultados de triagem de mutantes em dois organismos-modelo são mostrados na Figura 2.1. A ilustração mostra os efeitos dos mutantes no desenvolvimento da flor na planta *Arabidopsis thaliana* e no desenvolvimento do micélio no fungo *Neurospora crassa* (um micélio é uma rede de células filamentosas chamadas hifas). A ilustração mostra que o desenvolvimento das propriedades em questão pode ser alterado em uma variedade de modos diferentes. Na planta, o número ou tipo de órgãos florais é alterado; no fungo, a taxa de crescimento e o número e tipo de ramificações são alterados de uma variedade de modos, cada um dos quais dá uma morfologia de colônia distintamente anormal. A esperança é que cada caso represente uma mutação em um membro diferente do conjunto de genes responsáveis por essa propriedade. Entretanto, ocorrem as mudanças genéticas que são mais complexas que as mudanças de um gene; e, além disso, um ambiente anormal também é capaz de mudar o aspecto de um organismo. Assim, cada indivíduo deve ser testado para ver se ele produz descendentes na proporção apropriada que é diagnóstica de um mutante causado pela mutação de *um único gene*.

O enfoque genético para compreender uma propriedade biológica é descobrir que genes a controlam. Um enfoque para a descoberta do gene é isolar os mutantes e verificar cada um quanto aos padrões de herança monogênica (proporções específicas de expressão normal e mutante da propriedade nos descendentes).

Os padrões de herança monogênica são úteis para a descoberta do gene não só em genética experimental de organismos-modelo, mas também em genética aplicada. Um exemplo importante é encontrado em genética humana. Muitos distúrbios humanos, tais como fibrose cística e doença de Tay-Sachs, são herdados como um único gene mutante. Após um gene-chave ter sido definido desse modo, os geneticistas podem focá-lo ao nível do DNA e tentar decifrar o defeito

A análise genética começa com mutantes

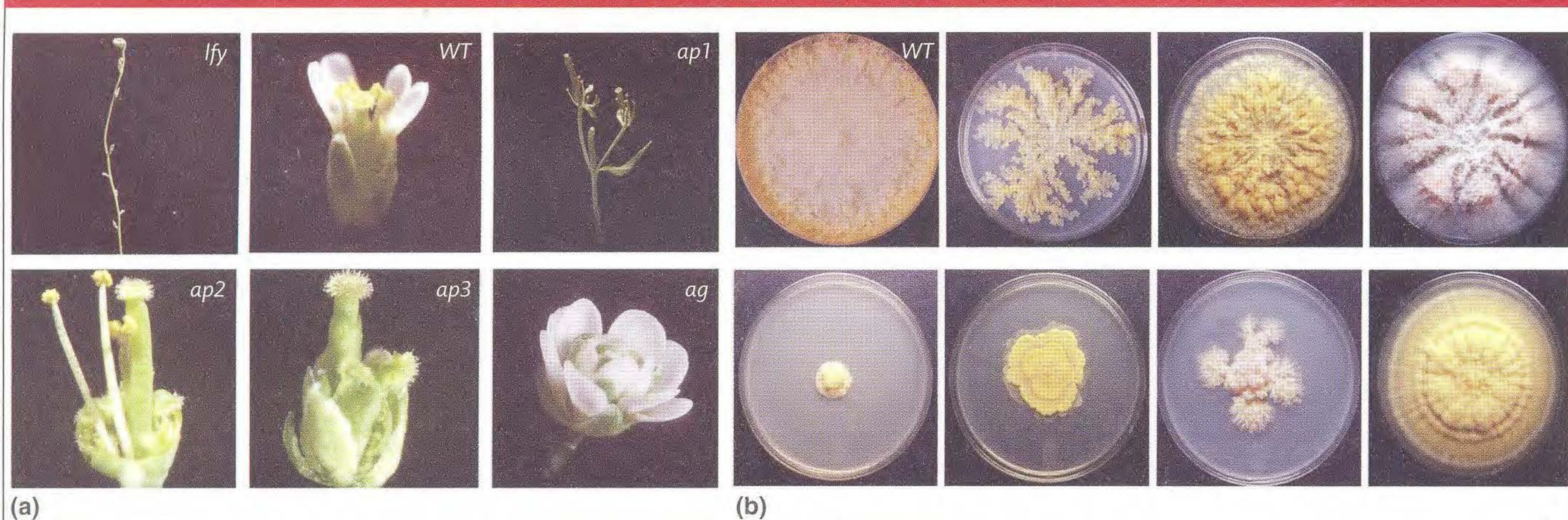


FIG. 2.1 Essas fotos mostram a gama de fenótipos mutantes típicos dos obtidos na dissecção genética de propriedades biológicas. Esses casos são da dissecção do desenvolvimento floral em *Arabidopsis thaliana* (a) e crescimento de hifas em *Neurospora crassa*, um fungo (b). WT = tipo selvagem. [(a) George Haughn; (b) Anthony Griffiths/Olivera Gavric.]

celular básico subjacente à doença, possivelmente levando a novas terapias. Na agricultura, esses mesmos tipos de padrões de herança monogênica levaram à descoberta de mutações que conferem alguma característica desejável, tal como resistência a uma doença ou melhor conteúdo de nutrientes. Essas mutações benéficas foram sucessivamente incorporadas em linhagens comerciais de plantas e animais.

As regras de herança monogênica foram originalmente elucidadas nos anos 1860 pelo monge Gregor Mendel, que trabalhou em um monastério na cidade de Brno, hoje parte da República Tcheca. A análise de Mendel é o protótipo do enfoque experimental da descoberta monogênica ainda usado hoje. De fato, Mendel foi o primeiro a descobrir qualquer gene! Mendel não sabia o que os genes eram, como eles influenciavam as propriedades biológicas, ou como eles eram herdados a nível celular. Hoje sabemos que os genes funcionam através de proteínas, um tópico ao qual voltaremos em capítulos posteriores. Também sabemos que os padrões de herança monogênica são produzidos porque os genes são partes dos cromossomos, e os cromossomos são repartidos muito precisamente ao longo das gerações. Assim, este capítulo começa com uma breve revisão dos genes e cromossomos como um prelúdio ao estudo de seus padrões de herança.

► 2.1 Genes e Cromossomos

Um conjunto único e completo de informação genética (DNA) de um organismo é chamado de seu **genoma**. A organização geral dos genomas em tipos diferentes de organismos e vírus

foi mostrado na Figura 1.14. Nosso foco no presente capítulo, e em todos os capítulos, até atingirmos o Capítulo 5, é nos genomas de espécies eucarióticas. Nos eucariontes, a maior parte do DNA de um genoma é encontrada no núcleo de cada célula. Esse DNA nuclear é dividido em unidades chamadas **cromossomos** (Figura 2.2). O conjunto de cromossomos presentes nos organismos da mesma espécie tem um número característico de cromossomos e aspecto. Um exemplo é visto na Figura 2.3, que mostra os cromossomos de uma célula do cervo indiano chamado muntjac. Essa ilustração revela algumas características interessantes dos cromossomos. A parte inferior mostra os cromossomos do núcleo, espalhados como resultado da quebra da membrana nuclear. Os cromossomos foram corados com sondas moleculares fluorescentes chamadas corantes cromossômicos, e, nessa preparação, as sondas foram escolhidas de modo que cromossomos idênticos tivessem a mesma cor. Essa coloração revela que o total de seis cromossomos é, na verdade, dois grupos de três: um par de vermelhos, um par de verdes e um par de cromossomos violeta. A presença desses pares aponta para uma característica importante do material genético nuclear da maioria dos animais e plantas: que eles são **diplóides**, significando isso que eles contêm dois genomas completos e, portanto, dois conjuntos de cromossomos idênticos. O número de cromossomos no conjunto genômico básico é chamado de **número haplóide** (chamado de n), que, para o muntjac, é 3. Assim, para o muntjac, o estado diplóide é designado $2n = 6$. Os seres humanos também são diplóides, mas, nesse caso, $n = 23$ e $2n = 46$.

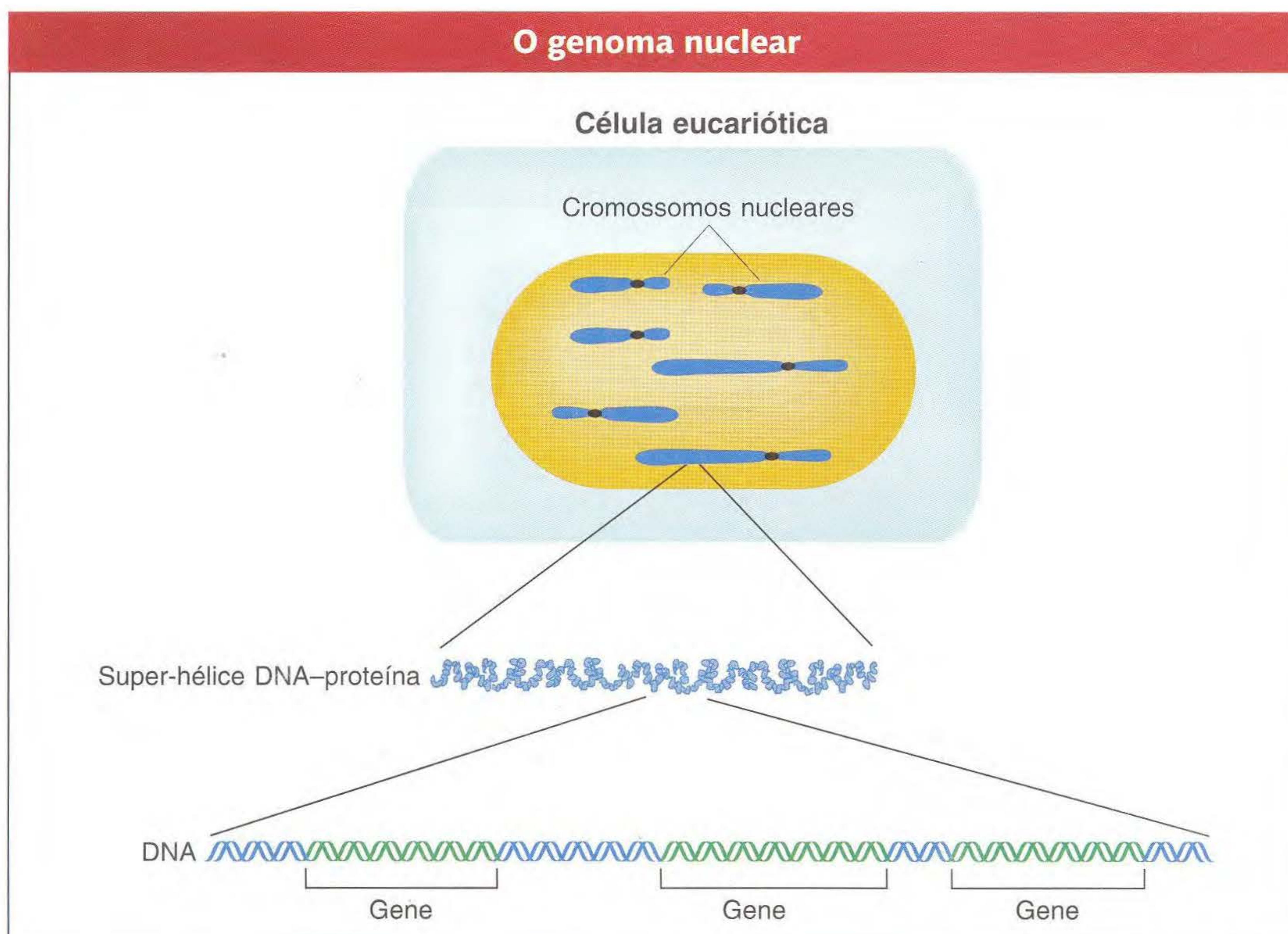


FIG. 2.2 O genoma nuclear é composto de um número espécie-específico de cromossomos. Uma região cromossômica foi expandida para mostrar o arranjo dos genes.

Um genoma diplóide visualizado

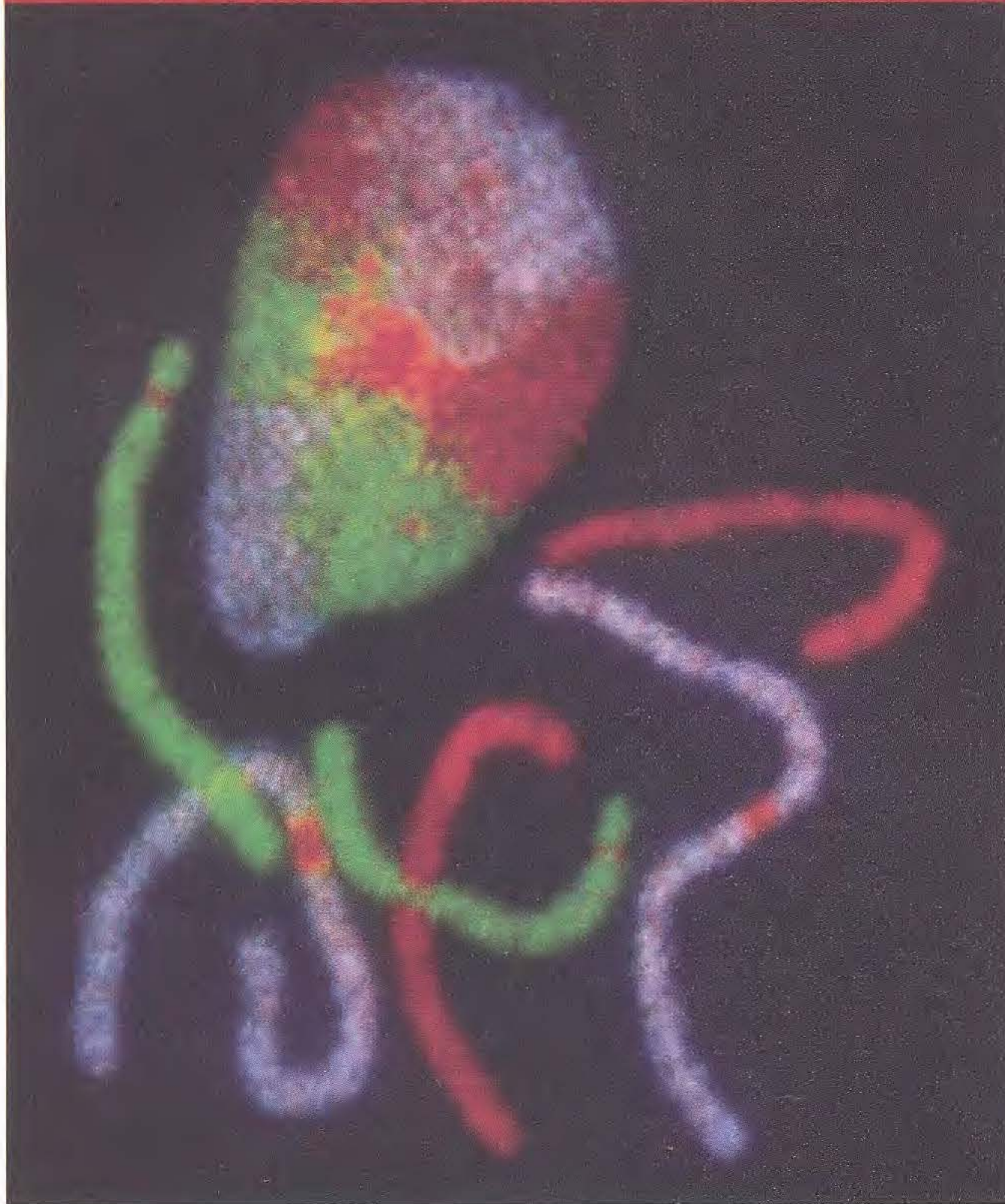


FIG. 2.3 O genoma nuclear nas células de uma fêmea do cervo indiano, um tipo de pequeno cervo ($2n = 6$). Os seis cromossomos visíveis são de uma célula obtida no processo de divisão nuclear. Os três pares de cromossomos foram corados com sondas de DNA específicas de cromossomos, cada uma marcada com um corante diferente (coloração de cromossomo). Um núcleo derivado de outra célula está no estágio entre as divisões. [Foto fornecida por Fengtang Yang and Malcolm Ferguson-Smith da Cambridge University. Apareceu como capa da *Chromosome Research*, vol. 6, no. 3, April 1998.]

Em um núcleo diplóide, os dois membros de um par de cromossomos são chamados de **cromossomos homólogos** ou, às vezes, apenas **homólogos**. As seqüências de DNA de membros de um par de homólogos geralmente são iguais, muito embora pequenas variações de seqüência em geral estejam presentes, o que é a base da variação genética dentro de uma espécie, o tipo de variação que nos permite distinguir uma da outra. Como os cromossomos homólogos são idênticos, eles levam os mesmos genes nas mesmas posições relativas. Assim, nos diplóides, cada gene está presente como um **par de genes**. Entretanto, note, na Figura 2.3, que, embora o núcleo em uma célula do corpo contenha pares de cromossomos, eles não estão fisicamente pareados no sentido de estarem próximos um do outro. Os cromossomos no núcleo rompido mostrados na parte inferior da imagem revelam falta de pareamento. Note também que a parte superior da imagem mostra um núcleo intacto de outra célula, e aqui, novamente, os cromossomos claramente não estão pareados. Por exemplo, os membros do par violeta estão em extremidades opostas ao núcleo. O pareamento físico dos homólogos ocorre na divisão nuclear conhecida como meiose, como veremos depois.

Muitas espécies, formando a maior parte da biomassa eucariótica no planeta, normalmente não são diplóides, mas **haplóides**. Em outras palavras, o núcleo contém apenas um conjunto de cromossomos. Os exemplos mais familiares de organismos haplóides são os fungos (incluindo o mofo, leveduras e cogumelos) e algas. Os haplóides tiveram papéis importantes na pesquisa genética como organismos-modelo porque são facilmente manipulados geneticamente, e, assim, freqüentemente voltaremos a eles neste livro.

Cada cromossomo contém uma molécula de DNA. Isso pode ser mostrado separando moléculas de DNA de um núcleo por tamanho usando eletroforese em gel, porque o número de bandas de DNA observadas após a eletroforese é igual ao número haplóide de cromossomos. Entretanto, os cálculos simples da quantidade de DNA por célula mostram que o tamanho de uma molécula de DNA em um cromossomo é muito maior que o tamanho do cromossomo. Por exemplo, o genoma humano tem cerca de 1 metro de DNA no total, mas está contido em 23 cromossomos, cada um medido em uma escala de milionésimos de um metro. Claramente, o DNA está compactado muito eficientemente em um cromossomo. Isso é obtido envolvendo o DNA ao redor de estruturas moleculares chamadas de **nucleossomos** (Figura 2.4). Cada nucleossomo é composto de um octâmero de proteínas chamadas *histonas*. A cadeia de DNA-nucleossomo é ainda mais helicoidizada e super-helicoidizada; o estado de um cromossomo eucariótico pode ser representado como na Figura 2.5. Essa ilustração mostra outro componente cromossômico chamado de arcabouço, que ajuda a organizar a estrutura tridimensional de um cromossomo. O DNA e nucleossomos associados constituem a **cromatina**, o arcabouço dos cromossomos.

A cromatina varia em compactação ao longo de um cromossomo. Conseqüentemente, os corantes que reagem com o DNA produzem intensidades diferentes de coloração ao longo de um cromossomo. A cromatina densa geralmente é encontrada ao redor do **centrômero**, uma parte de um cromossomo em geral visível como uma constrição. (Veremos que o centrômero tem um papel em levar os cromossomos para as células descendentes na divisão celular.) A cromatina densa é chamada de **heterocromatina**, e a menos densa é chamada de **eucromatina** (Figura 2.6). Além de tornarem os cromossomos compactos, as histonas que se ligam ao DNA podem afetar a regulação gênica, como veremos no Capítulo 11.

Outra característica do conjunto cromossômico visível na Figura 2.6 é o **organizador nucleolar**, uma região única (geralmente em um cromossomo específico) que contém várias repetições de genes codificando RNA ribossomal. Um corpo esférico contendo RNA ribossomal, o **nucléolo**, é geralmente visível ligado ao organizador nucleolar. Para completar, devemos mencionar duas outras estruturas cromossômicas: telômeros e bandas cromossômicas. Os **telômeros** são as pontas dos cromossomos. Embora os telômeros em geral não tenham características visíveis, eles são interessantes porque usam um único mecanismo de replicação diferente do resto do cromossomo. Quando corados, os cromossomos de algumas espécies em geral apresentam **bandas cromossômicas** transversas características, das quais veremos casos em capítulos posteriores. O centrômero, a heterocromatina,

DNA cromossômico enrolado ao redor de histonas

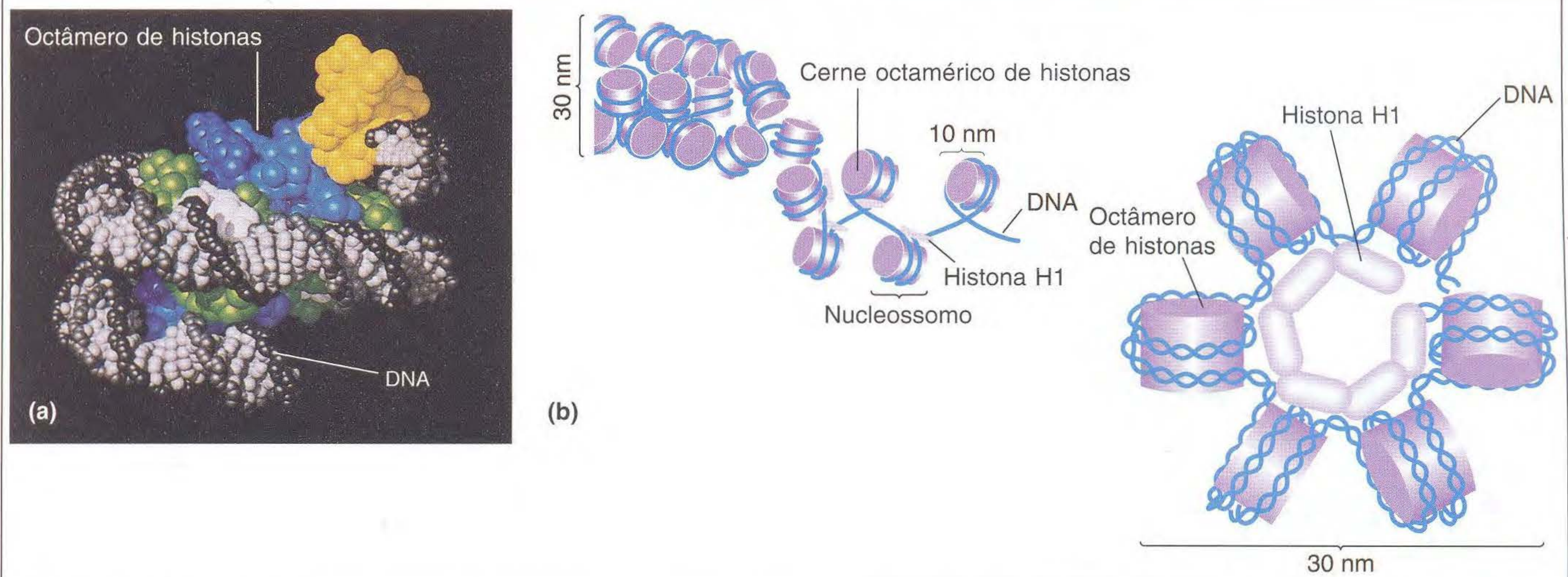


FIG. 2.4 (a) Um modelo de um nucleossomo mostra o DNA envolvendo um octâmero de histonas. (b) Visão lateral e terminal da cadeia helicoidal de nucleossomos, diâmetro de 30 nm, mostrando os octâmeros de histona como discos púrpura. Uma histona adicional chamada H1, que não é parte do octâmero, corre pelo centro da hélice agindo como um estabilizador. [(a) Alan Wolffe and Van Moudrianakis; (b) H. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, and J. Darnell, *Molecular Cell Biology*, 3rd ed. Copyright 1995 by Scientific American Books.]

Condensação cromossômica por super-helicoidização

Estrutura 3D
do cromossomo

W.W. ANIMAÇÃO COMPUTADORIZADA

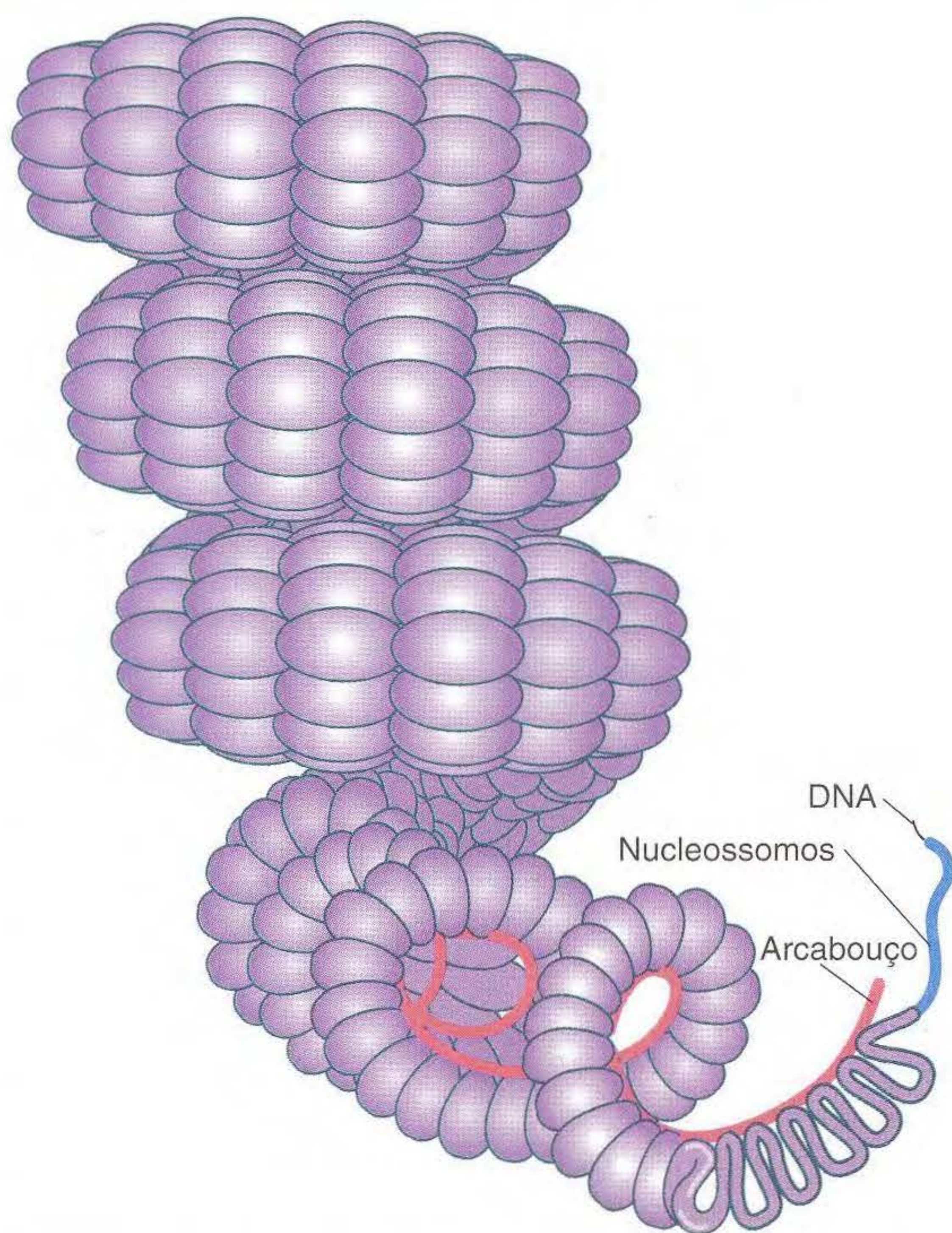


FIG. 2.5 O modelo mostra um cromossomo super-helicoidizado na divisão celular. As voltas são tão densamente compactadas que só suas pontas são visíveis. Em uma ponta, a estrutura é parcialmente deseloidizada para mostrar os componentes.

Alguns marcos do cromossomo 2 de tomate

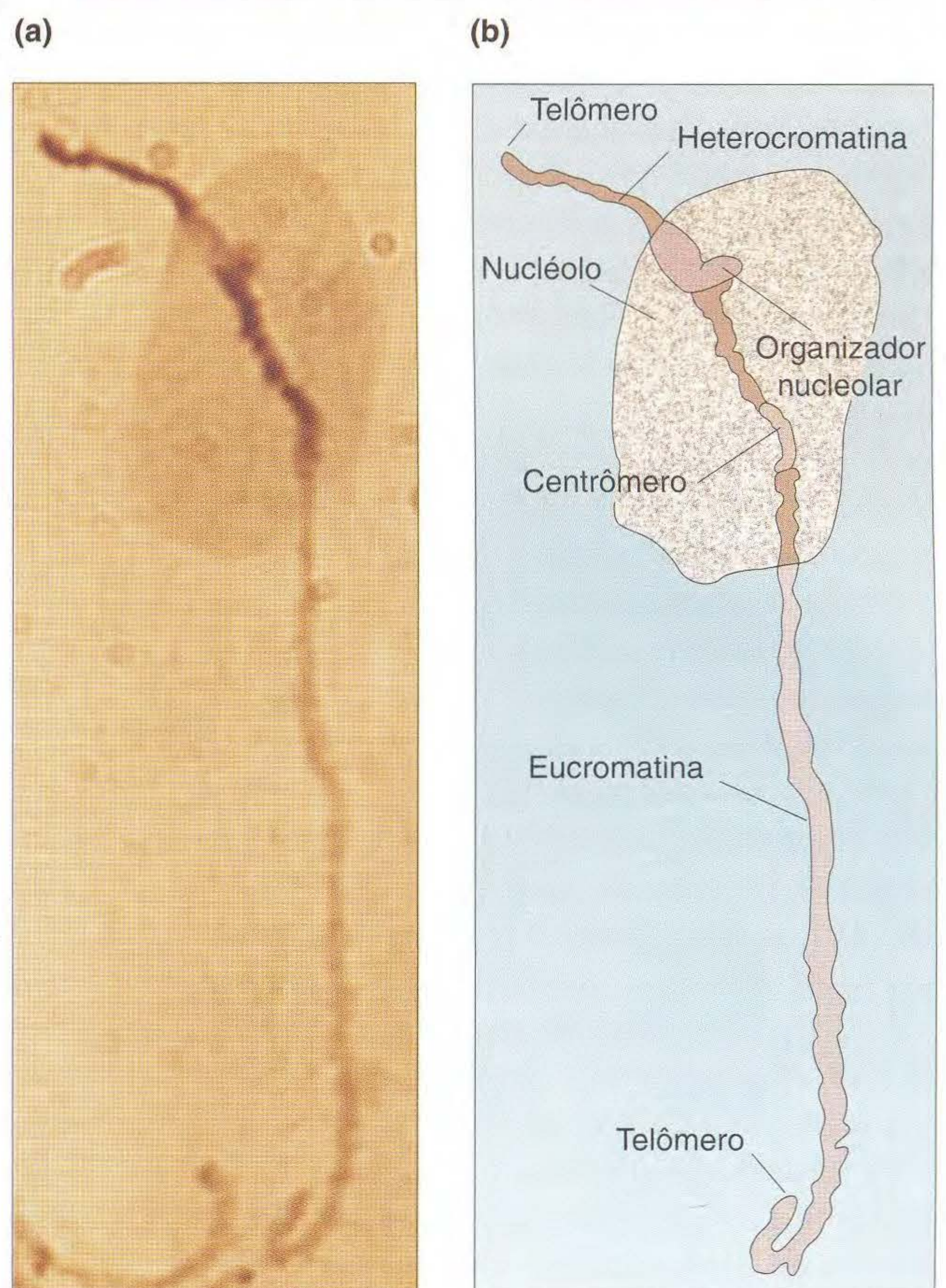


FIG. 2.6 (a) Foto; (b) interpretação. [Foto fornecida por Peter Moens; de P. Moens and L. Butler. "The Genetic Location of the Centromere of Chromosome 2 in the Tomato." *Can. J. Genet. Cytol.* 5, 1963, 364-370.]

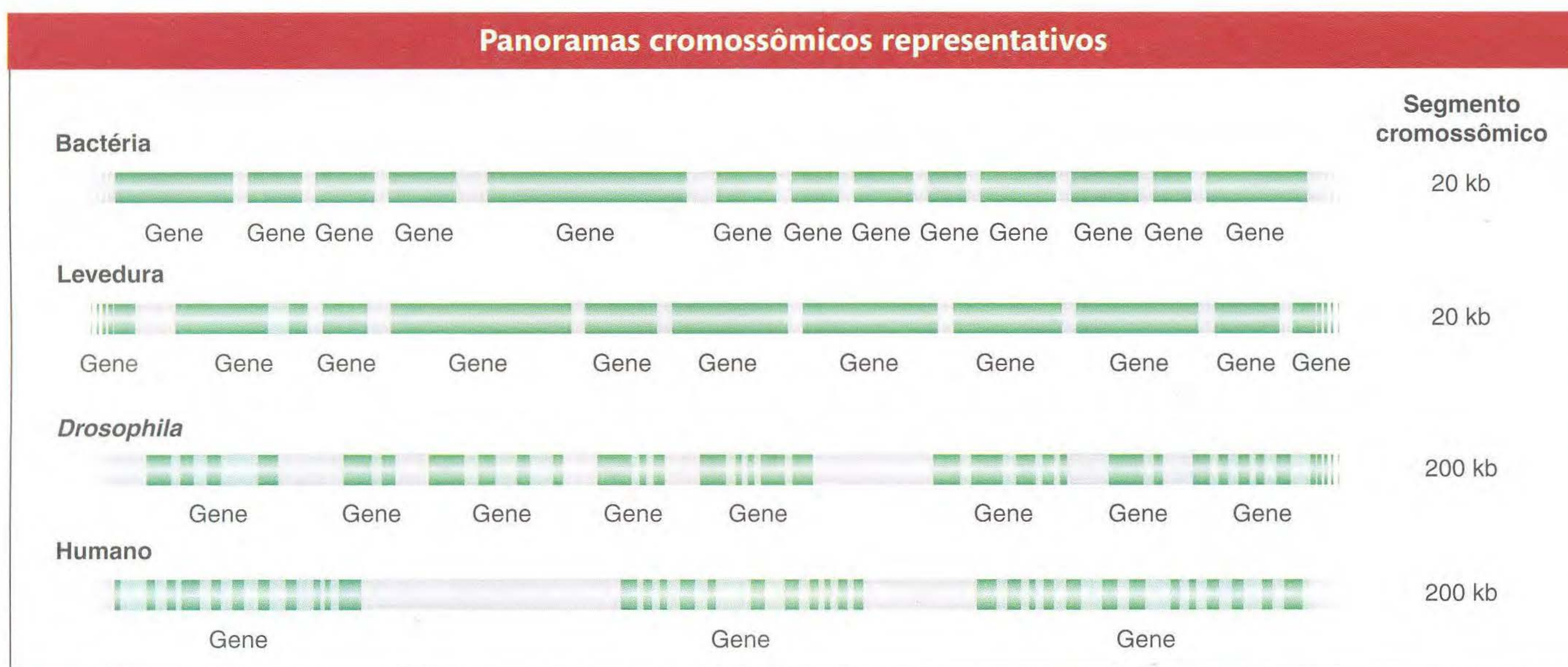


FIG. 2.7 Os genomas de quatro espécies diferentes têm topografias gênicas muito diferentes. Verde-claro, íntrons; verde-escuro, éxons; branco, regiões entre seqüências codificantes (incluindo regiões reguladoras mais DNA “espaçador”). As duas ilustrações de cima e as duas de baixo estão em escalas diferentes.

o organizador nucleolar e as bandas cromossômicas são marcos cromossômicos importantes para a análise genética.

Em todos os organismos, os genes são segmentos transcritos dispostos ao longo do DNA de um cromossomo. Entretanto, existe uma considerável variação entre espécies no número de genes e no “panorama” cromossômico geral. O número de genes varia de cerca de 6.000, na levedura *Saccharomyces cerevisiae*, até aproximadamente 25.000, no *Homo sapiens*. As regiões não-transcritas entre os genes são variáveis entre as espécies em tamanho e conteúdo de DNA. Um achado surpreendente em uma pesquisa recente é que, em muitas espécies eucarióticas, uma grande parte desse espaço intergênico resultou do acúmulo de um tipo “móvel” de DNA chamado de **elemento de transposição**. A maioria dos eucariontes tem uma grande quantidade de DNA cuja origem é claramente de elementos de transposição, embora alguns, tais como o fungo *Neurospora*, tenham muito poucos. Como o genoma pode tolerar esse acúmulo e permanecer funcional nós não sabemos, mas os elementos de transposição são conhecidos como tendo um efeito no rearranjo do genoma no curso da evolução, como veremos no Capítulo 14. Outra fonte de variação entre as espécies (e outra surpresa que emerge das pesquisas moleculares) é a prevalência e tamanho dos **íntrons**, as regiões não-codificantes que interrompem os segmentos codificantes de um gene. O tamanho da região codificante de um gene (geralmente mais ou menos sinônima do tamanho de seu mRNA) é marcadamente constante entre as espécies, variando geralmente entre 1 e 3 quilobases (kb). Entretanto, a presença de grandes números de íntrons pode tornar enorme o tamanho aparente do gene. Um caso mais extremo é o gene humano para a proteína distofina, que é defeituosa na doença muscular distrofia: esse gene tem 78 íntrons, que fazem com que o gene total tenha 2.500 kb de tamanho, muito embora o tamanho do mRNA esteja na faixa normal. Uma amostra generalizada

dos panoramas cromossômicos é mostrada na Figura 2.7, e um exemplo específico do genoma humano, mostrado na Figura 2.8, ilustra o arranjo dos íntrons em alguns genes reais.

► 2.2 Padrões de Herança Monogênica

A inserção dos genes dentro dos cromossomos, revista na seção anterior, é crítica para a compreensão do modo pelo qual os genes são transmitidos pelas gerações. Concorde, vejamos o modo pelo qual podemos identificar os padrões de herança mostrados por um único gene. Lembre que a primeira etapa na dissecação genética é obter variantes que diferem na propriedade em questão. Com a suposição de que obtivemos uma coleção de mutantes relevantes, a questão seguinte é se cada uma das mutações é herdada como um gene único.

A lei de Mendel da segregação igual

A primeira análise de herança monogênica como uma via para a descoberta do gene foi feita por Gregor Mendel. Foi dele a análise que seguiremos como exemplo. Mendel escolheu a ervilha-de-cheiro, *Pisum sativum*, como seu organismo de pesquisa. A escolha do organismo para qualquer pesquisa biológica é crucial, e a escolha de Mendel provou ser boa porque as ervilhas são facilmente cultivadas e reproduzidas. Note, entretanto, que Mendel não fez uma caçada de mutantes em ervilhas; de fato, ele fez uso dos mutantes que já haviam sido encontrados por outros na horticultura. Além disso, o trabalho de Mendel difere da pesquisa da maioria dos geneticistas de hoje em dia porque não foi uma dissecação genética; ele não estava interessado nas propriedades das ervilhas em si, mas no modo pelo qual as unidades hereditárias que influenciam essas propriedades foram herdadas de geração a geração. Entretanto, as leis da herança deduzidas

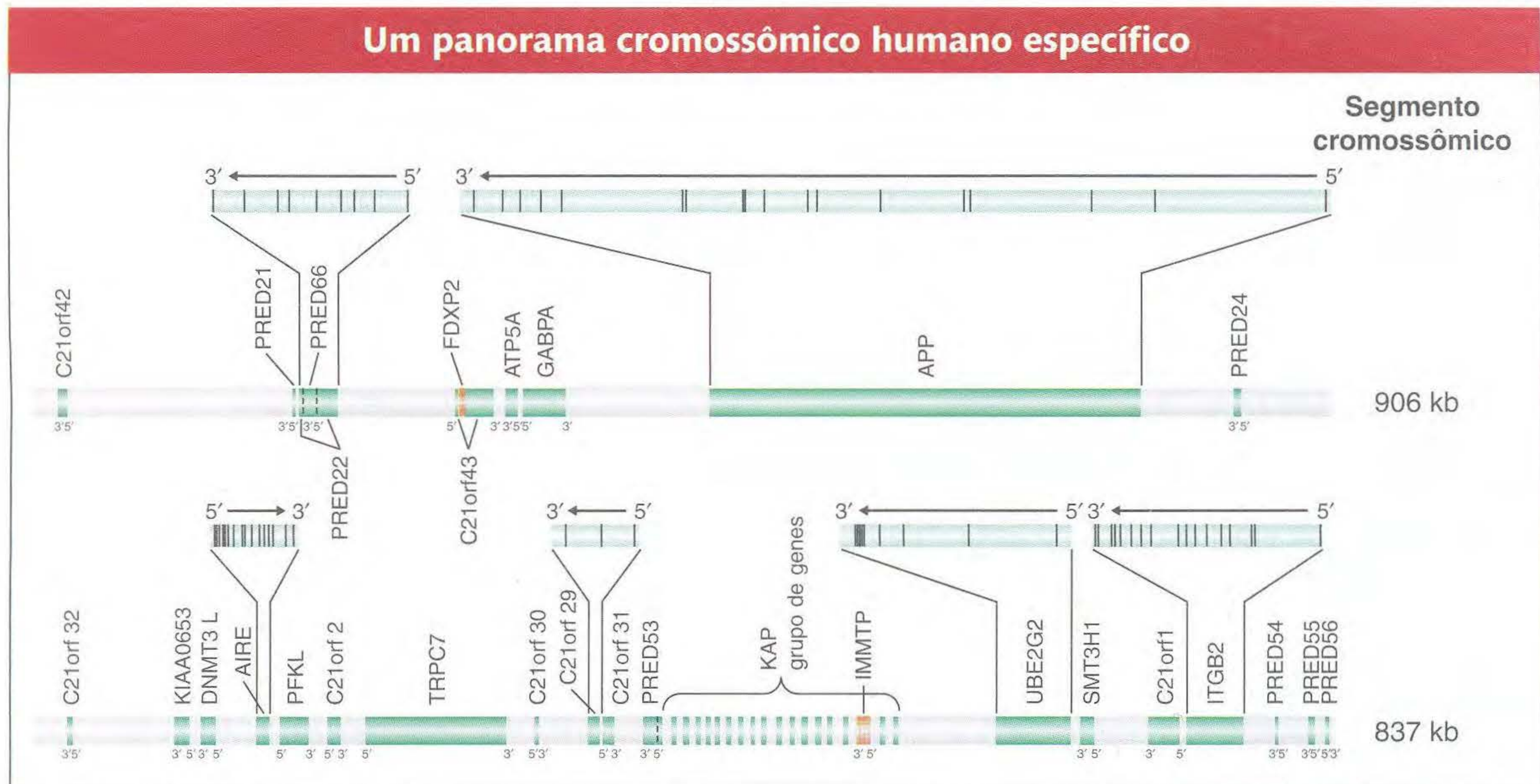


FIG. 2.8 Regiões transcritas de genes (verde) em dois segmentos do cromossomo 21, com base na sequência completa desse cromossomo. (Dois genes, FDXP2 e IMMTP, estão em laranja para distinguir-se dos genes vizinhos.) Alguns genes estão expandidos para mostrar éxons (barras pretas) e introns (em verde-claro). As marcações verticais são nomes de genes (alguns com função conhecida e outros não). As marcações 5' e 3' mostram a direção da transcrição dos genes. [De M. Hattori et al., *Nature* 405, 2000, 311-319.]

por Mendel são exatamente as que usamos hoje em dia na genética moderna para identificar padrões monogênicos de herança.

Mendel escolheu investigar a herança de sete propriedades dessa espécie escolhida de ervilhas: cor da ervilha, forma da ervilha, cor da vagem, forma da vagem, cor da flor, altura da planta e posição do broto da planta. Em genética, os termos **caráter** e **característica** são usados mais ou menos como sinônimos de **propriedade**. Para cada uma dessas

sete características, ele obteve de seu fornecedor horticultural duas linhagens que apresentavam aspectos distintos e contrastantes. Hoje em dia, diríamos que, para cada característica, ele estudou dois **fenótipos** contrastantes. Um fenótipo pode ser definido como uma *forma de uma característica*. Esses fenótipos contrastantes são ilustrados na Figura 2.9. Seus resultados foram substancialmente os mesmos para ambas as características, e, assim, podemos usar uma característica, cor da semente de ervilha, como uma ilustra-

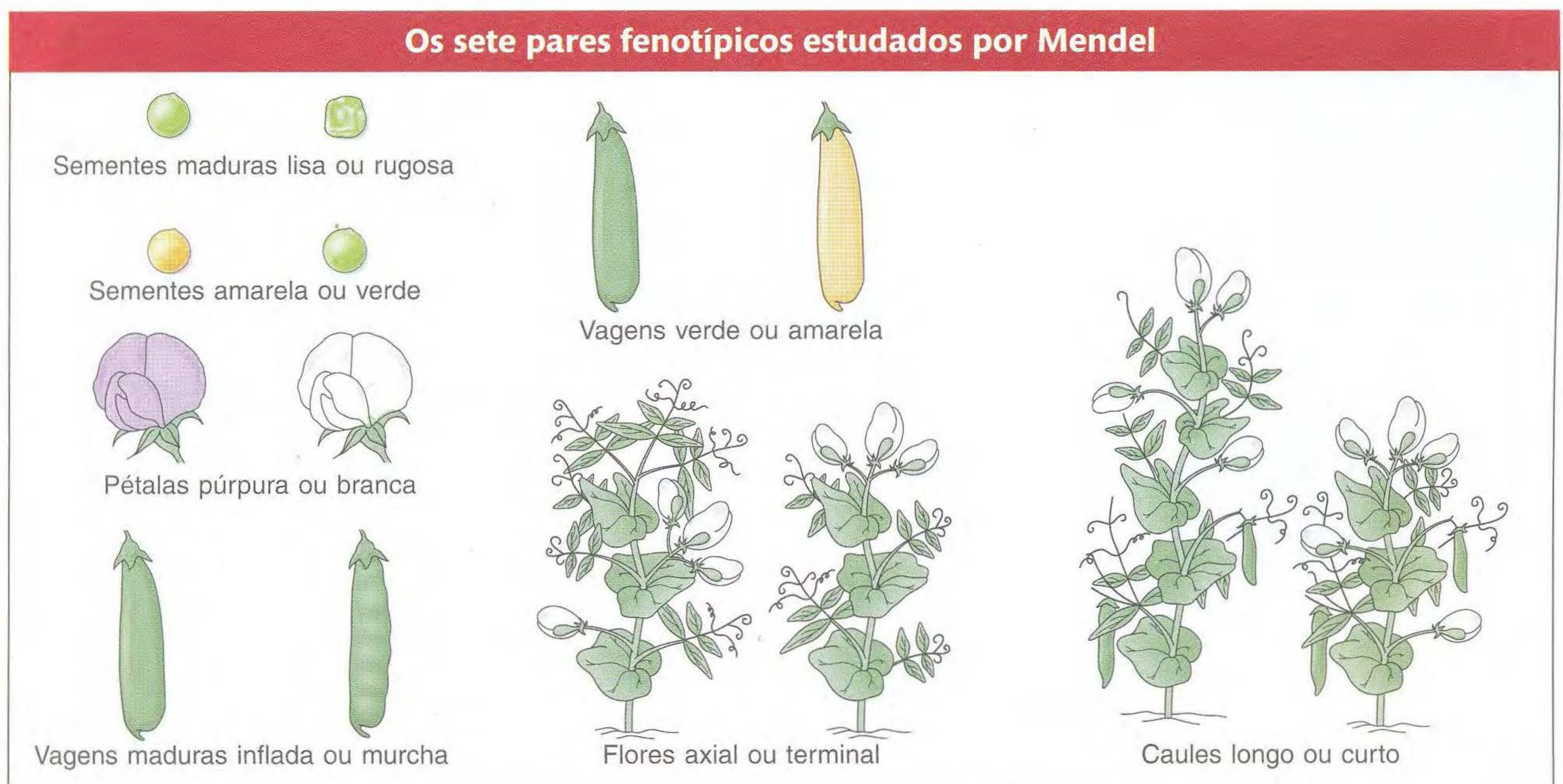


FIG. 2.9 Para cada característica, Mendel estudou dois fenótipos contrastantes. [De S. Singer and H. Hilgard, *The Biology of People*. Copyright 1978 by W.H. Freeman and Company.]

ção. Todas as linhagens usadas por Mendel eram **linhagens puras**, significando que, para o fenótipo em questão, toda a prole produzida pelos cruzamentos entre os membros dessa linhagem eram idênticos. Por exemplo, dentro da linhagem de sementes amarelas, toda a prole de qualquer cruzamento era de sementes amarelas.

A análise de Mendel da hereditariedade em ervilhas fez um amplo uso desses cruzamentos. Para fazer um cruzamento em plantas tais como ervilha, o pólen é simplesmente transferido das anteras de uma planta para os estigmas de outra. Um tipo especial de cruzamento é a **autofecundação**, que é feita deixando o pólen de uma flor cair sobre seus próprios estigmas. O cruzamento e a autofecundação são ilustrados na Figura 2.10. O primeiro cruzamento feito por Mendel foi de plantas da linhagem de sementes amarelas com plantas da linhagem de sementes verdes. Essas linhagens constituíam a **geração parental**, abreviada como P. Em *Pisum sativum*, a cor da semente (a ervilha) é determinada por sua própria composição genética; assim, as ervilhas resultantes de um cruzamento são efetivamente a prole e podem ser convenientemente classificadas pelo fenótipo sem a necessidade de cultivá-las em plantas. A prole de ervilhas do cruzamento entre as linhagens puras era toda de amarelas, não importa que genitor (amarelo ou verde) tivesse sido usado como masculino ou feminino. Essa geração prole é chamada de **primeira geração filial**, ou F_1 . Assim, os resultados desses dois cruzamentos recíprocos foram os seguintes, onde X representa um cruzamento:

fêmea de linhagem amarela \times macho de linhagem verde
 \rightarrow ervilhas de F_1 todas amarelas
 fêmea de linhagem verde \times macho de linhagem amarela
 \rightarrow ervilhas de F_1 todas amarelas

Os resultados observados nos descendentes de ambos os cruzamentos recíprocos foram iguais, e, assim, nós os tratamos como um cruzamento. Mendel cultivou as ervilhas de F_1 em plantas, e autofecundou ou entrecruzou as plantas F_1 resultantes para obter a **segunda geração filial**, ou F_2 . A F_2 era composta de 6.022 ervilhas amarelas e 2.001 ervilhas verdes. Mendel notou que esse resultado era muito próximo de uma proporção matemática exata de três quartos de amarelas e um quarto de verdes. Curiosamente, o fenótipo verde, que tinha desaparecido na F_1 , tinha reaparecido em um quarto dos indivíduos F_2 , mostrando que os determinantes genéticos para verde deveriam estar presentes na F_1 amarela, embora não expressos.

Em seguida Mendel *autofecundou individualmente* plantas crescidas das sementes de F_2 . As plantas cultivadas das sementes verdes F_2 , quando autofecundadas, foram encontradas gerando apenas ervilhas verdes. Entretanto, as plantas cultivadas de sementes amarelas de F_2 , quando autofecundadas individualmente, foram encontradas de dois tipos: um terço delas eram puras para sementes amarelas, mas dois terços delas deram uma prole de três quartos de sementes amarelas e um quarto de sementes verdes, do mesmo modo que as plantas de F_1 .

Outro cruzamento informativo que Mendel fez foi da F_1 com qualquer planta de semente verde; aqui, a prole mostrou as proporções de metade amarela e metade verde. Esses dois tipos de cruzamentos, a F_1 autofecundada e o cruzamento da F_1 com qualquer planta de sementes verdes, ambos deram prole amarela e verde, mas em proporções diferentes. Essas duas proporções são representadas na Figura 2.11. Note que as proporções são vistas apenas quando as ervilhas de várias vagens são combinadas.

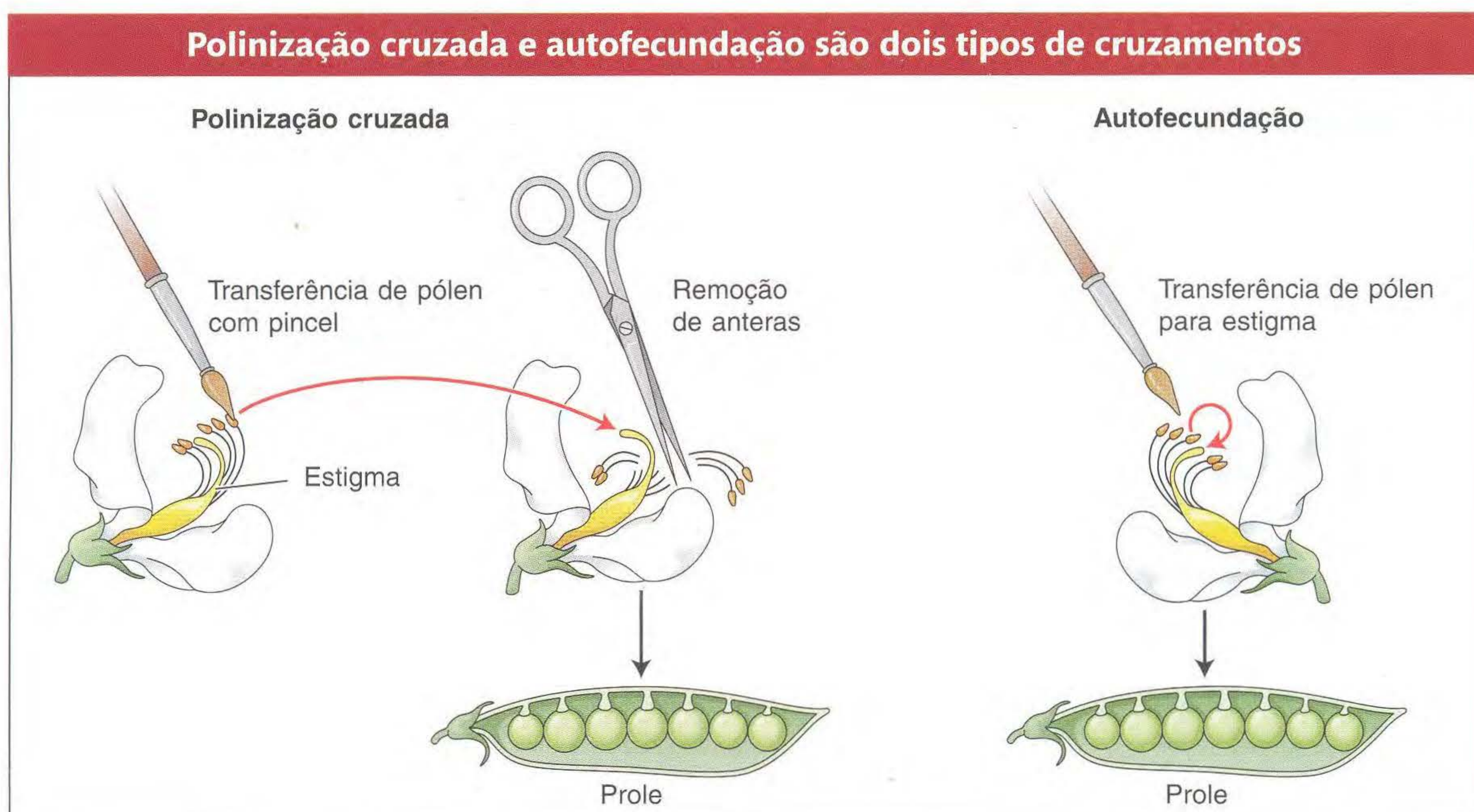


FIG. 2.10 Em um cruzamento de ervilhas (esquerda), o pólen das anteras de uma planta é transferido para o estigma de outra. Na autofecundação (direita), o pólen é transferido das anteras para os estigmas da mesma planta.

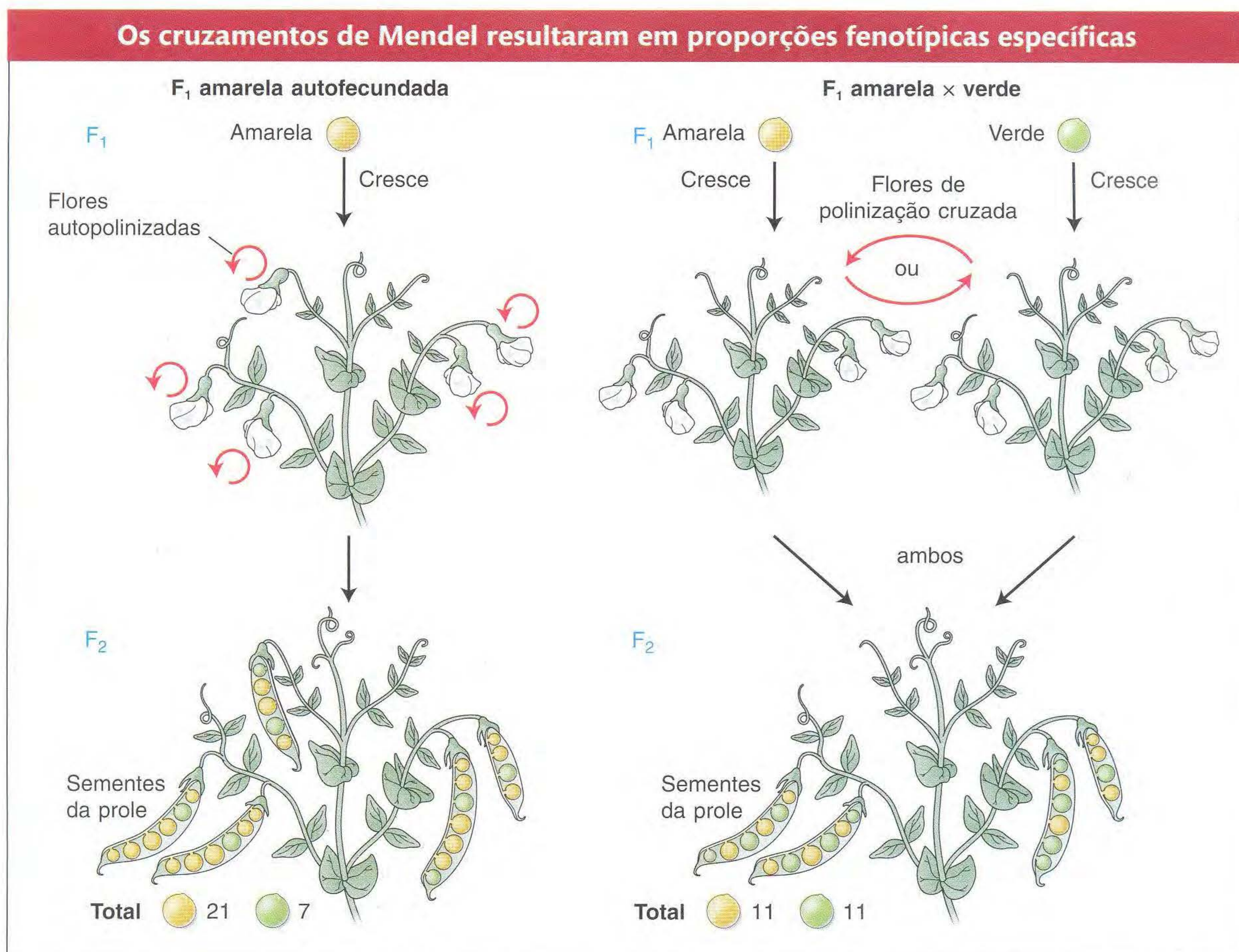


FIG. 2.11 Mendel obteve uma proporção fenotípica de 3:1 em sua autopolinização de F₁ (esquerda) e uma proporção fenotípica de 1:1 em seu cruzamento de F₁ amarela com verde (direita). Os tamanhos das amostras são arbitrários.

As proporções de 3:1 e 1:1 encontradas para cor de ervilhas também foram encontradas para cruzamentos comparáveis das outras seis características que Mendel estudou. Os números reais dessas proporções 3:1 para essas características são mostrados no Quadro 2.1. Inicialmente, o significado dessas proporções matemáticas precisas e repetíveis não devem ter sido claras para Mendel, mas ele foi capaz de criar um modelo brilhante que não só respondia por todos os resultados, como também representou o nascimento histórico da ciência genética. O modelo de Mendel para o exemplo da cor da ervilha, traduzido em termos modernos, era o seguinte.

1. Um fator hereditário chamado **gene** era necessário para produzir a cor da ervilha.
2. Cada planta tem um par desse tipo de gene.
3. O gene existe em duas formas chamadas de **alelos**. Se o gene é foneticamente chamado de gene “wye” (em inglês), então os dois alelos podem ser representados por Y (representando o fenótipo *yellow*, amarelo) e y (representando o fenótipo verde).
4. Uma planta pode ser ou Y/Y, y/y ou Y/y. A barra mostra que os alelos estão em um par.

Quadro 2.1 Resultados dos Cruzamentos Mendelianos nos quais os Genitores Diferem em uma Característica

Fenótipo parental	F ₁	F ₂	Proporção de F ₂
1. lisa × sementes rugosas	Todas lisas	5.474 lisas; 1.850 rugosas	2,96:1
2. amarela × sementes verdes	Todas amarelas	6.022 amarelas; 2.001 verdes	3,01:1
3. púrpura × pétalas brancas	Todas púrpura	705 púrpura; 224 brancas	3,15:1
4. inflada × vagem murcha	Todas infladas	882 infladas; 299 murchas	2,95:1
5. verde × vagens amarelas	Todas verdes	428 verdes; 152 amarelas	2,82:1
6. axial × flores terminais	Todas axiais	651 axiais; 207 terminais	3,14:1
7. longo × caules curtos	Todos longos	787 longos; 277 curtos	2,84:1

5. Na planta Y/y , o alelo Y domina, e, assim, o fenótipo será amarelo. Logo, o fenótipo da planta Y/y define o alelo Y como **dominante** e o alelo y como **recessivo**.
6. Na meiose, os membros de um par de genes separam-se igualmente em ovócitos e espermatozoides. Essa separação igual tornou-se conhecida como **Primeira Lei de Mendel** ou como **lei da segregação igual**.
7. Assim, um único gameta contém apenas um membro de cada par.
8. Na fertilização, os gametas se fundem aleatoriamente, independentemente de qual dos alelos ele leva.

Aqui, introduzimos alguma terminologia. Um ovócito fertilizado, a primeira célula que se desenvolve no indivíduo da prole, é chamado de **zigoto**. Uma planta com um par de alelos idênticos é chamada de **homozigota**, uma planta na qual os alelos do par diferem é chamada de **heterozigota**. Às vezes, um heterozigoto para um gene é chamado de **monoíbrido**. Um indivíduo pode ser classificado como **homozigoto dominante** (tal como Y/Y), **heterozigoto** (Y/y) ou **homozigoto recessivo** (y/y). Geralmente, em genética, as combinações alélicas subjacentes aos fenótipos são chamadas de **genótipos**. Logo, Y/Y , Y/y e y/y são todos genótipos.

A Figura 2.12 mostra como os postulados de Mendel explicam as proporções da prole ilustradas na Figura 2.11. As linhagens “puras” são homozigotas, seja Y/Y ou y/y . Logo, cada linhagem produz apenas gametas Y ou apenas gametas y e, assim, dão só homozigotos. Quando cruzadas umas com

as outras, as linhagens Y/Y e y/y produzem uma geração F_1 composta de todos os indivíduos heterozigotos (Y/y). Como Y é dominante, todos os indivíduos da F_1 são de fenótipo amarelo. A autofecundação dos indivíduos F_1 pode ser imaginada como um cruzamento do tipo $Y/y \times Y/y$. Um cruzamento do tipo $Y/y \times Y/y$ às vezes é chamado de **cruzamento monoíbrido**. A segregação igual dos alelos Y e y nos heterozigotos de F_1 resulta em gametas, tanto masculinos quanto femininos, metade dos quais são Y e metade y . Os gametas masculinos e femininos fundem-se aleatoriamente na fertilização, com os resultados mostrados na grade da Figura 2.12. A composição de F_2 é três quartos de sementes amarelas e um quarto de verdes, uma proporção de 3:1. A proporção de um quarto das sementes F_2 que são verdes dão sempre verdes, como esperado pelo genótipo y/y . Entretanto, as sementes amarelas F_2 (totalizando três quartos) são de dois genótipos: dois terços delas são claramente heterozigotas Y/y e um terço, homozigotas dominantes Y/Y . Logo, vemos que a proporção fenotípica 3:1 subjacente na F_2 é uma proporção genotípica 1:2:1, assim:

$$\left. \begin{array}{l} \frac{1}{4} Y/Y \text{ amarela} \\ \frac{2}{4} Y/y \text{ amarelas} \\ \frac{1}{4} y/y \text{ verde} \end{array} \right\} \frac{3}{4} \text{ amarelas (Y/-)}$$

A representação geral de um indivíduo expressando o alelo dominante é $Y/-$; a barra representando um espaço que pode

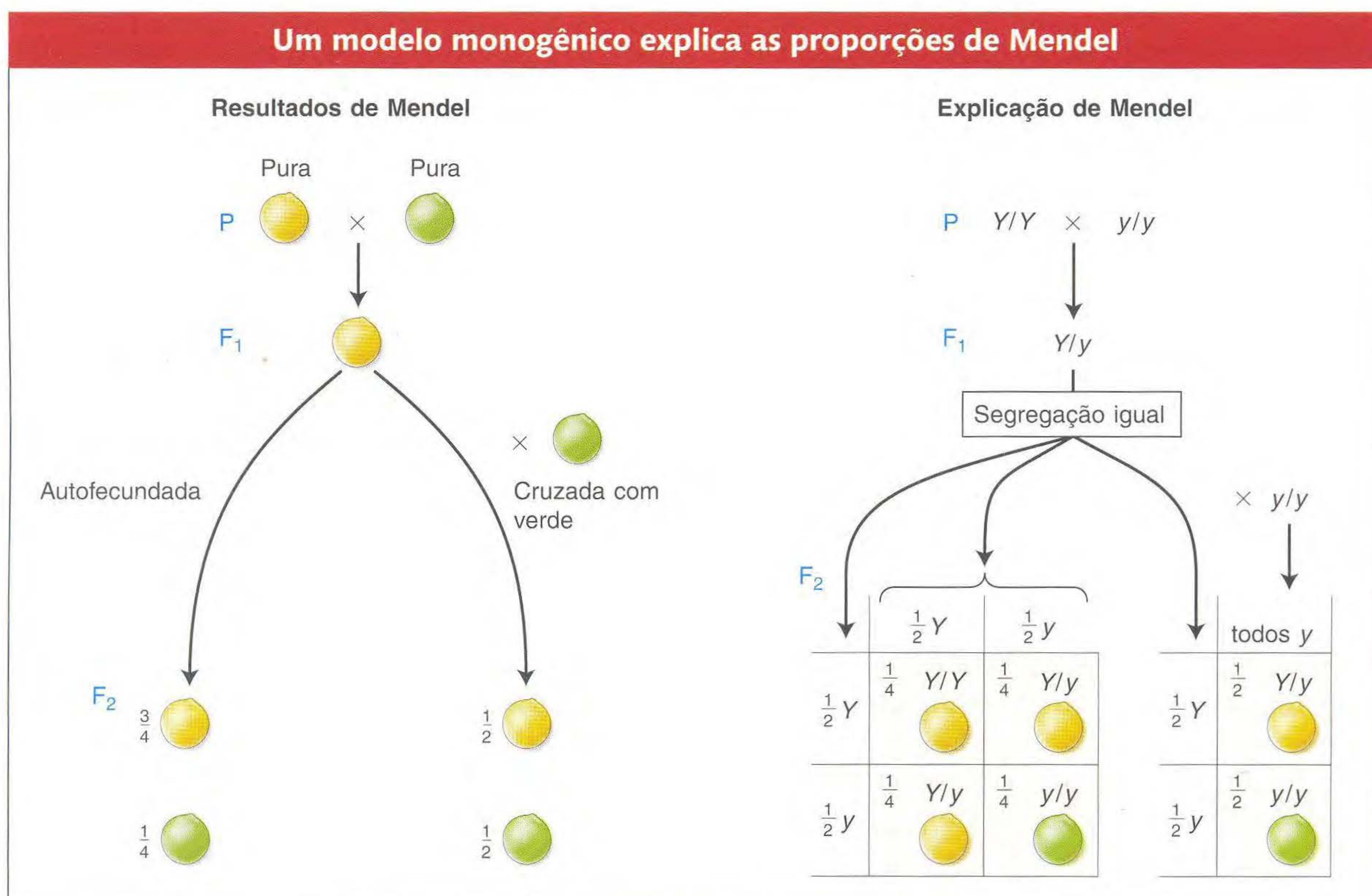


FIG. 2.12 Os resultados de Mendel (esquerda) são explicados por um modelo monogênico (direita) que postula uma segregação igual dos membros de um par de genes para gametas.

ser preenchido com outro Y ou um y . Note que a segregação igual é detectável apenas na meiose de um heterozigoto. Assim, Y/y produz metade de gametas Y e metade de gametas y . Embora a segregação igual também esteja ocorrendo nos homozigotos, nem a segregação $\frac{1}{2}Y:\frac{1}{2}Y$ nem a segregação $\frac{1}{2}y:\frac{1}{2}y$ é significativa ou detectável a nível genético.

Agora também podemos explicar os resultados do cruzamento entre as plantas crescidas de sementes F_1 amarelas (Y/y) e plantas crescidas de sementes verdes (y/y). Nesse caso, a segregação igual nos heterozigotos amarelos de F_1 dá gametas com uma proporção $\frac{1}{2}Y:\frac{1}{2}y$. O genitor y/y só pode fazer gametas y ; logo, o fenótipo da prole só depende de quais alelos eles herdaram do genitor Y/y . Assim, a proporção gamética $\frac{1}{2}Y:\frac{1}{2}y$ do heterozigoto é convertida em uma proporção genotípica $\frac{1}{2}Y/y:\frac{1}{2}y/y$, o que corresponde a uma proporção fenotípica de 1:1 de plantas de semente amarela para plantas de semente verde. Isso é ilustrado no painel à direita na Figura 2.12.

Note que, na definição dos pares de alelos subjacentes a esses fenótipos, Mendel identificou um gene que afeta radicalmente a cor da ervilha. Essa identificação não era seu interesse principal, mas nós podemos ver que descobrir padrões monogênicos de herança é um processo de descoberta do gene, identificando genes que influenciam uma propriedade biológica.

As proporções 1:1, 3:1 e 1:2:1 são todas diagnósticas de herança monogênica e baseadas na segregação igual em um heterozigoto.

A pesquisa de Mendel, na metade do século XIX, não foi notada pela comunidade científica internacional até que observações similares foram independentemente publicadas por vários outros pesquisadores em 1900. Logo, a pesquisa em muitas espécies de plantas, animais, fungos e algas mostraram que a lei de Mendel da segregação igual era aplicável a todos os eucariontes e, em todos os casos, era baseada nas segregações cromossômicas que ocorrem na meiose, tópico ao qual voltaremos na seção seguinte.

2.3 A Base Cromossômica dos Padrões de Herança Monogênica

A visão de Mendel da segregação igual era que os membros de um par de genes segregam-se igualmente na formação de gametas. Ele não conhecia os eventos subcelulares que ocorrem quando as células se dividem no curso da formação de gametas. Hoje em dia, compreendemos que pares de genes estão situados em pares de cromossomos, e são os membros de um par de cromossomos que na verdade segregam, levando os genes consigo. Os membros de um par de genes são segregados como uma consequência inevitável.

Quando as células se dividem, também se dividem o núcleo e seus principais conteúdos, os cromossomos. Para entender a segregação gênica, devemos primeiro compreender e contrastar os dois tipos de divisões nucleares que ocorrem nas

Estágios do ciclo celular assexual

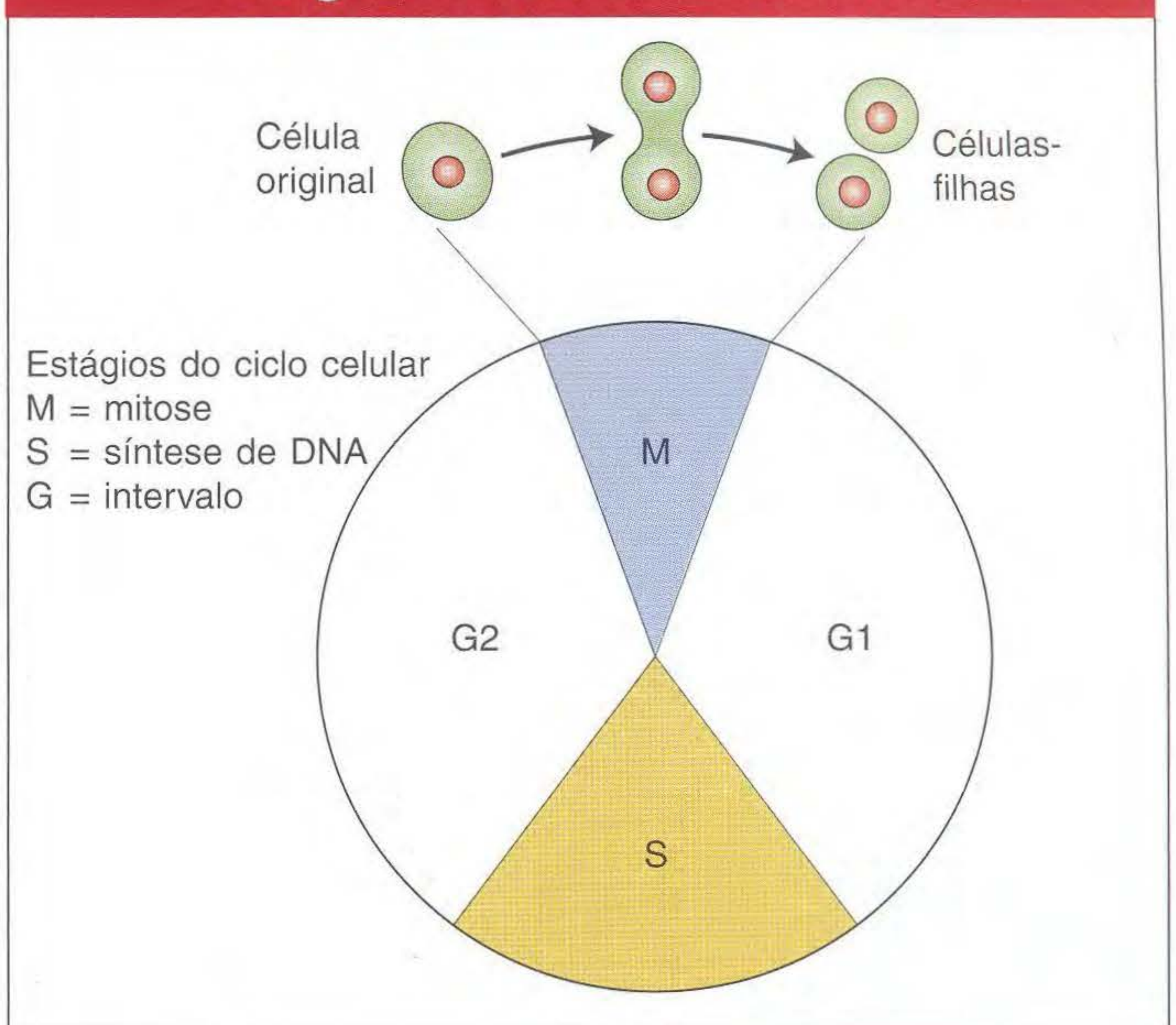


FIG. 2.13

células eucarióticas. Quando as células somáticas se dividem para aumentar seu número, a divisão nuclear acompanhante é chamada de **mitose**, um estágio programado de todos os ciclos de divisão celular (Figura 2.13). A mitose pode ocorrer em células diplóides ou haplóides. Como resultado, a célula genitora torna-se duas. Assim,

$$\begin{aligned} \text{ou } 2n &\longrightarrow 2n + 2n \\ \text{ou } n &\longrightarrow n + n \end{aligned}$$

Além disso, a maioria dos eucariontes tem um ciclo sexual, e, nesses organismos, células diplóides especializadas, chamadas **meiócitos**, são separadas para produzir células sexuais, tais como espermatozóides e ovócitos em plantas e animais ou esporos sexuais em fungos ou algas. Ocorrem duas divisões celulares sequenciais, e as duas divisões nucleares que as acompanham são chamadas de **meiose**. Como existem duas divisões, são produzidas quatro células. A meiose ocorre apenas em células diplóides, e as células resultantes (espermatozóides e ovócitos em animais e plantas) são haplóides. Assim, o resultado final da meiose é

$$2n \longrightarrow n + n + n + n$$

A localização dos meiócitos nos ciclos vitais de animais, plantas e fungos é mostrada na Figura 2.14.

As características genéticas básicas da mitose e meiose estão resumidas na Figura 2.15. Para facilitar a comparação, ambos os processos são mostrados em uma célula diplóide. Observe, novamente, que a mitose ocorre em uma divisão celular, e que as duas células-filhas têm o mesmo conteúdo genômico que a célula-mãe (genitora). O primeiro processo importante a notar é uma replicação cromossômica pré-mitótica. Ao nível do DNA, esse estágio é a síntese, ou fase S (veja Figura 2.13), no qual o DNA é

Divisão celular em ciclos de vida comuns

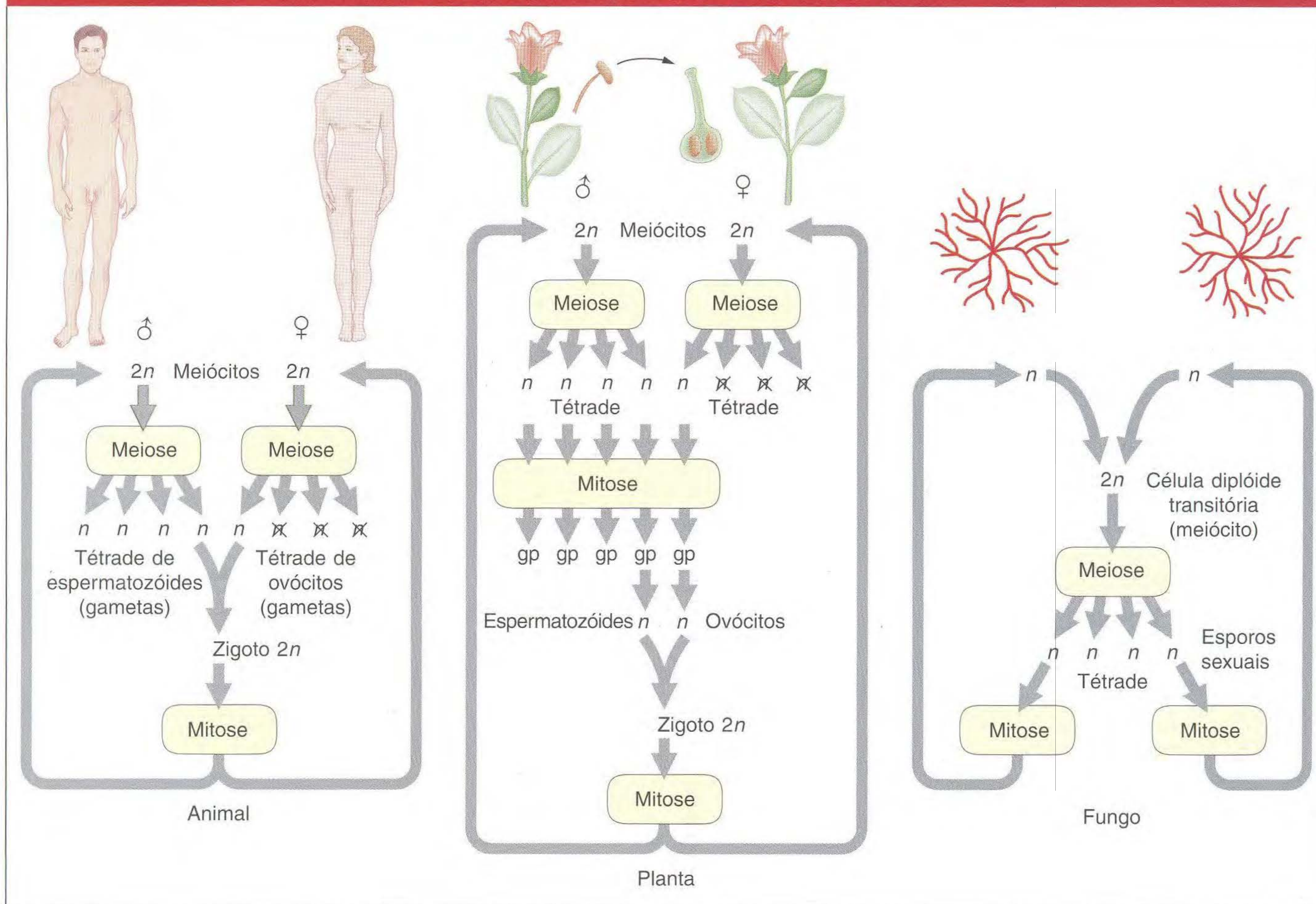
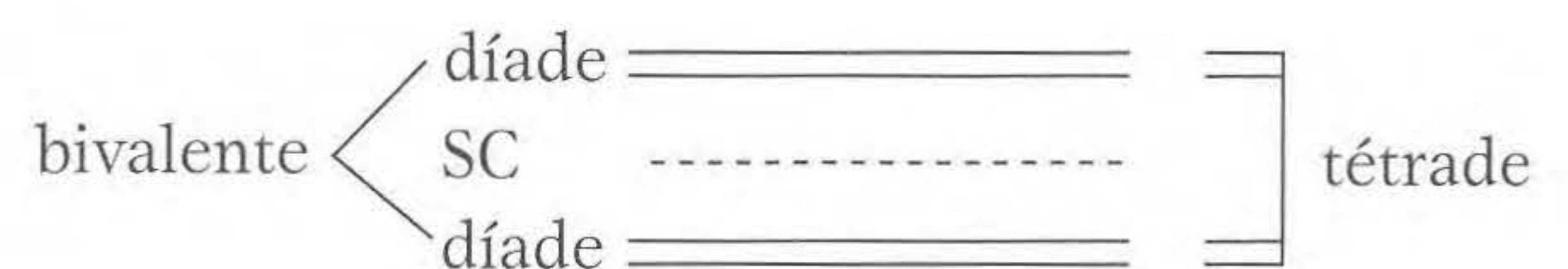


FIG. 2.14 Os ciclos de vida de humanos, plantas e fungos, mostrando os pontos nos quais a mitose e a meiose ocorrem. Note que, nas fêmeas humanas e de muitas plantas, três células da tétrade meiótica abortam. A abreviação n indica uma célula haplóide, $2n$ uma célula diplóide; gp representa "gametófito", o nome da pequena estrutura composta de células haplóides que produzirão gametas. Em muitas plantas tais como o milho, um núcleo do gametófito masculino funde-se com dois núcleos do gametófito feminino, dando origem a uma célula triplóide ($3n$), que então se replica para formar o endosperma, um tecido nutritivo que circunda o embrião (que é derivado do zigoto $2n$).

replicado. A replicação produz pares de **cromátides-irmãs** idênticas, que se tornam visíveis no começo da mitose. Quando uma célula se divide, cada membro de um par de cromátides-irmãs é levado para cada célula-filha, onde assume o papel de um cromossomo comum. Assim, cada célula-filha tem o mesmo conteúdo cromossômico que a célula original.

Antes da meiose, como na mitose, a replicação cromossômica ocorre para formar cromátides-irmãs, que se tornam visíveis na meiose. O centrômero parece não se dividir nesse estágio, enquanto o faz na mitose. Também, em contraste com a mitose, os pares de homólogos das cromátides-irmãs unem-se para formar um feixe de quatro cromátides homólogos. Essa união de pares de homólogos é chamada de *sinapse*, e é baseada nas propriedades de montagem macromolecular chamadas de complexo sinaptonêmico (SC), que corre pelo centro do par (Figura 2.16). Os cromossomos-irmãos replicantes em conjunto são chamados de **díade** (da palavra grega para dois). A unidade que compreende o par de díades em sinapse é chamada de um **bivalente**. As quatro cromátides

que constituem um bivalente são chamadas de **tétrade** (do grego para quatro), para indicar que existem quatro unidades no feixe.



(Uma nota entre parênteses. O processo de *crossing over* ocorre nesse estágio de tétrade. O crossing over muda as combinações de alelos de vários genes diferentes, mas não afeta diretamente os padrões de herança monogênica; portanto, adiaremos sua cobertura detalhada até o Capítulo 4. No momento, devemos notar que, além de sua função de combinar alelos, o crossing over também é conhecido como sendo um evento crucial que deve ocorrer para que exista a segregação cromossômica apropriada na primeira divisão meiótica.)

Os bivalentes de todos os cromossomos movem-se para o equador da célula, e, quando a célula se divide, uma díade

vai para cada nova célula, levada pelas fibras do fuso ligadas aos centrômeros. Na segunda divisão celular da meiose, os centrômeros dividem-se, e cada membro de uma díade (cada membro de um par de cromátides) move-se para uma célula-filha. Assim, embora o processo comece com o mesmo conteúdo genômico que a mitose, as duas segregações sucessivas resultam em quatro células haplóides. Cada uma das quatro células haplóides que constituem os quatro **produtos da meiose** contém um membro de uma tétrade, e, por isso, o grupo de quatro células às vezes também é chamado de uma tétrade. A meiose pode ser resumida do seguinte modo:

Início: \longrightarrow dois homólogos

Replicação: \longrightarrow duas díades

Pareamento: \longrightarrow tétrade

Primeira divisão: \longrightarrow uma díade para cada célula-filha

Segunda divisão: \longrightarrow uma cromátide para cada célula-filha

O comportamento dos cromossomos durante a meiose explica claramente a lei de Mendel da segregação igual. Considere um heterozigoto do tipo geral A/a . Podemos simplesmente seguir o resumo precedente considerando o que ocorre com os alelos desse gene:

Começo: um homólogo leva A e um leva a

Replicação: uma díade é AA e uma é aa

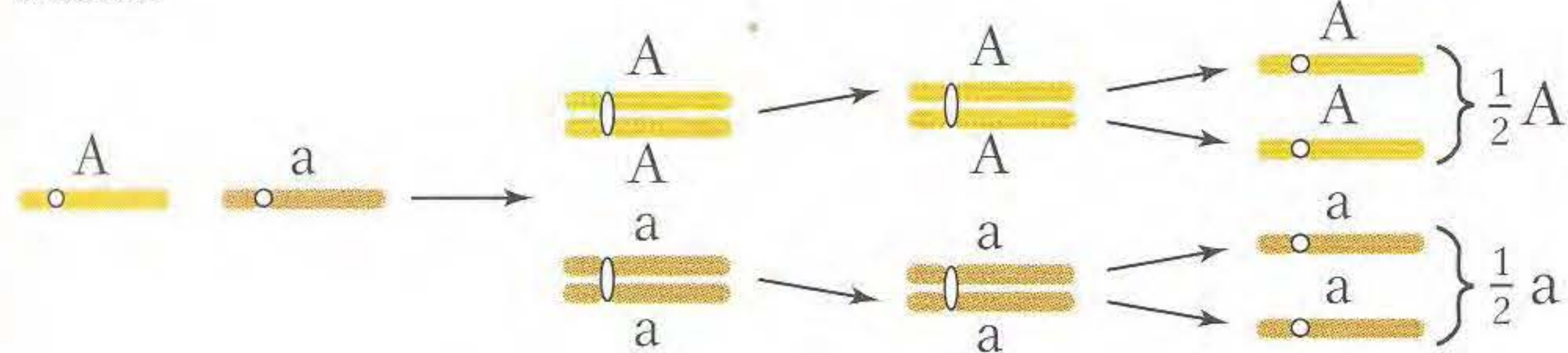
Pareamento: $A/A/a/a$

Produtos da primeira divisão: uma célula AA , a outra célula aa (o crossing over pode misturar esses tipos de produtos, mas a proporção geral não é alterada)

Produtos da segunda divisão: quatro células, duas do tipo A e duas do tipo a

Assim, os produtos da meiose de um meiócito heterozigoto A/a são $\frac{1}{2} A$ e $\frac{1}{2} a$, exatamente a proporção igual à necessária para explicar a Primeira Lei de Mendel.

Meiose



Note que, na presente discussão, enfocamos aspectos genéticos amplos da meiose, necessários para explicar a herança monogênica. Descrições mais completas dos estágios detalhados da mitose e meiose são apresentadas nos Boxes 2.1 e 2.2.

Herança monogênica em haplóides

Vimos que a base celular da lei da segregação igual é a segregação de cromossomos na primeira divisão da meiose. Na discussão até agora, a evidência da segregação igual de alelos nos meiócitos tanto de plantas quanto de animais é

indireta, com base na observação de que os cruzamentos apresentam as proporções apropriadas da prole esperada sob segregação igual. Reconhece que os gametas nesses estudos (tais como os de Mendel) devem ter surgido de *muitos diferentes* meiócitos. Entretanto, em alguns organismos haplóides, tais como várias espécies de fungos e algas, a segregação igual pode ser observada *diretamente* dentro de um meiócito *individual*. O motivo é que, nos ciclos desses organismos, os quatro produtos de uma única meiose são temporariamente mantidos juntos em um tipo de saco. O fermento de padeiro, *Saccharomyces cerevisiae*, fornece um bom exemplo (veja o boxe Organismo-modelo *Levedura* no Capítulo 11). Nos fungos, existem formas simples de sexos chamadas *tipos reprodutivos*. Em *S. cerevisiae*, os dois tipos reprodutivos são chamados MAT α e MAT a , determinados pelos alelos de um gene. (Note que o símbolo de um gene pode ter mais de uma letra — nesse caso, quatro.) Um cruzamento bem-sucedido deve ser entre as linhagens de tipos reprodutivos opostos, isto é, MAT α \times MAT a . Vejamos um cruzamento que inclui uma levedura mutante. As colônias de leveduras tipo selvagem normais são brancas, mas, ocasionalmente, surgem mutantes vermelhos devido a uma mutação em um gene na via bioquímica que produz adenina. Usemos o mutante vermelho para investigar a segregação igual em um único meiócito. Podemos chamar o alelo mutante r de *vermelho*. Que símbolo podemos usar para o alelo normal, ou tipo selvagem? Em genética experimental, o alelo tipo selvagem para qualquer gene em geral é designado pelo sinal $+$. Esse sinal é ligado a um superescrito do símbolo inventado para o alelo mutante. Assim, o alelo tipo selvagem nesse exemplo seria designado r^+ , mas um simples $+$ em geral é usado para abreviação. Para ver a segregação de um gene, o mutante vermelho é cruzado com o tipo selvagem. O cruzamento teria que estar entre tipos reprodutivos diferentes. Por exemplo, se o mutante vermelho tiver surgido em uma linhagem MAT α , o cruzamento seria

$$\text{MAT}\alpha \cdot r^+ \times \text{MAT}a \cdot r$$

Quando duas células de tipo reprodutivo oposto se fundem, é formada uma célula diplóide, e é essa célula que se torna um meiócito. No presente exemplo (ignorando o tipo reprodutivo para focar a mutação vermelha), o meiócito diplóide seria heterozigoto r^+/r . A replicação e segregação de r^+ e r daria uma tétrade de dois produtos meióticos (esporos) de genótipo r^+ e dois de r , todos contidos dentro de um saco membranoso chamado **asco**. Assim,

$$r^+/r \longrightarrow \left. \begin{array}{c} r^+ \\ r^+ \\ r \\ r \end{array} \right\} \text{tétrade no asco}$$

Os detalhes do processo são mostrados na Figura 2.17. Se forem isolados os quatro esporos de um asco (representando uma tétrade de cromátides) e usados para gerar quatro culturas de leveduras, então a segregação igual dentro de um meiócito é revelada diretamente como duas culturas brancas e duas vermelhas. Se analisarmos os esporos aleatórios de

muitos meiócitos, encontraremos cerca de 50% vermelhos e 50% brancos.

Note a simplicidade dos haplóides genéticos: um cruzamento requer a análise apenas de uma meiose; em contraste, um cruzamento diplóide requer a consideração da meiose tanto no genitor masculino quanto no feminino. Essa simplicidade é um motivo importante para usar haplóides como organismos-modelo. Outro motivo é que, nos haplóides, todos os alelos são expressos no fenótipo, pois não há mascaramento dos recessivos pelos alelos dominantes no outro homólogo.

A base molecular da segregação e expressão monogênica

A pesquisa moderna revelou processos moleculares que expandem nossa compreensão de muitos dos termos e conceitos que usamos para descrever padrões de herança monogênica.

Replicação O que está ocorrendo a nível molecular durante a formação de cromátides-irmãs, que ocorre como um prelúdio tanto da meiose quanto da mitose? Sabemos que o componente genômico primário de cada cromossomo é uma molécula de DNA. Essa molécula de DNA é replicada durante a fase S que precede tanto a mitose quanto a meiose. Como veremos no Capítulo 7, a replicação é um processo preciso, e, portanto, toda a informação genética é duplicada, seja tipo selvagem ou mutante. Por exemplo, se um mutante é o resultado de uma mudança em um único par de nucleotídeos, digamos, de GC (tipo selvagem) para AT (mutante), então, em um heterozigoto, a replicação será a seguinte:

homólogo GC → replicação → cromátide GC
cromátide GC

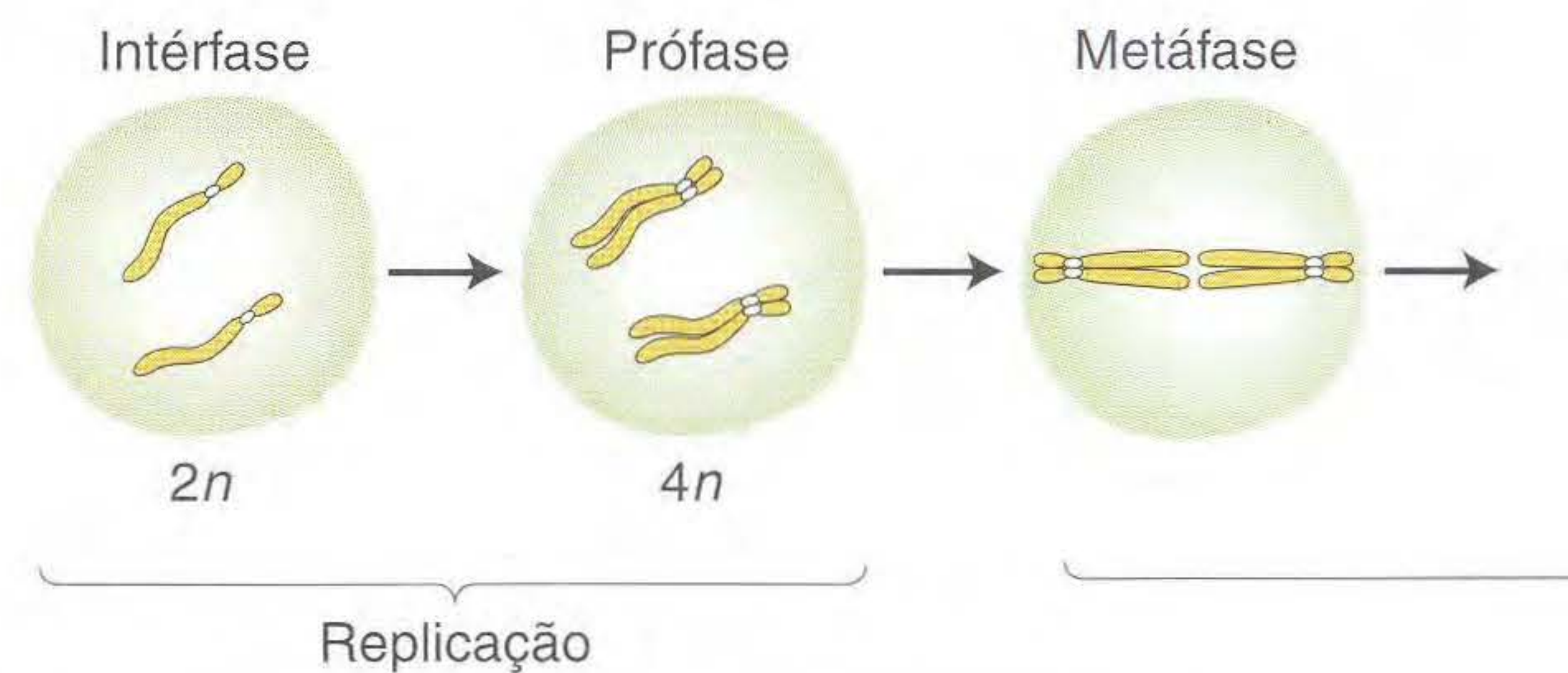
homólogo AT → replicação → cromátide AT
cromátide AT

A replicação do DNA antes da mitose em um haplóide e um diplóide é mostrada na Figura 2.18. Esse tipo de ilustração serve para nos lembrar que, em nossas considerações sobre os mecanismos de herança, essencialmente são as moléculas de DNA que estão sendo movidas nas células em divisão. A divisão nuclear visualizada ao nível de DNA molecular é mostrada na Figura 2.19.

Demonstração da segregação cromossômica ao nível molecular Interpretamos os padrões de herança monogênica fenotípica em relação à segregação de DNA cromossômico na meiose. Existe algum modo para mostrar diretamente a segregação do DNA? O método da força bruta seria seqüenciar os alelos (digamos, *A* e *a*) nos produtos meióticos, e o resultado seria que metade dos produtos teria uma seqüência *A* de DNA e metade teria a seqüência de DNA *a*. O modo mais simples é usar um *polimorfismo de comprimento de fragmentos de restrição (RFLP)*. No Capítulo 1, aprendemos que as enzimas de restrição são enzimas bacterianas que cortam o DNA em seqüências específicas de bases no genoma. As seqüên-

Estágios principais da meiose e mitose

Mitose



Meiose

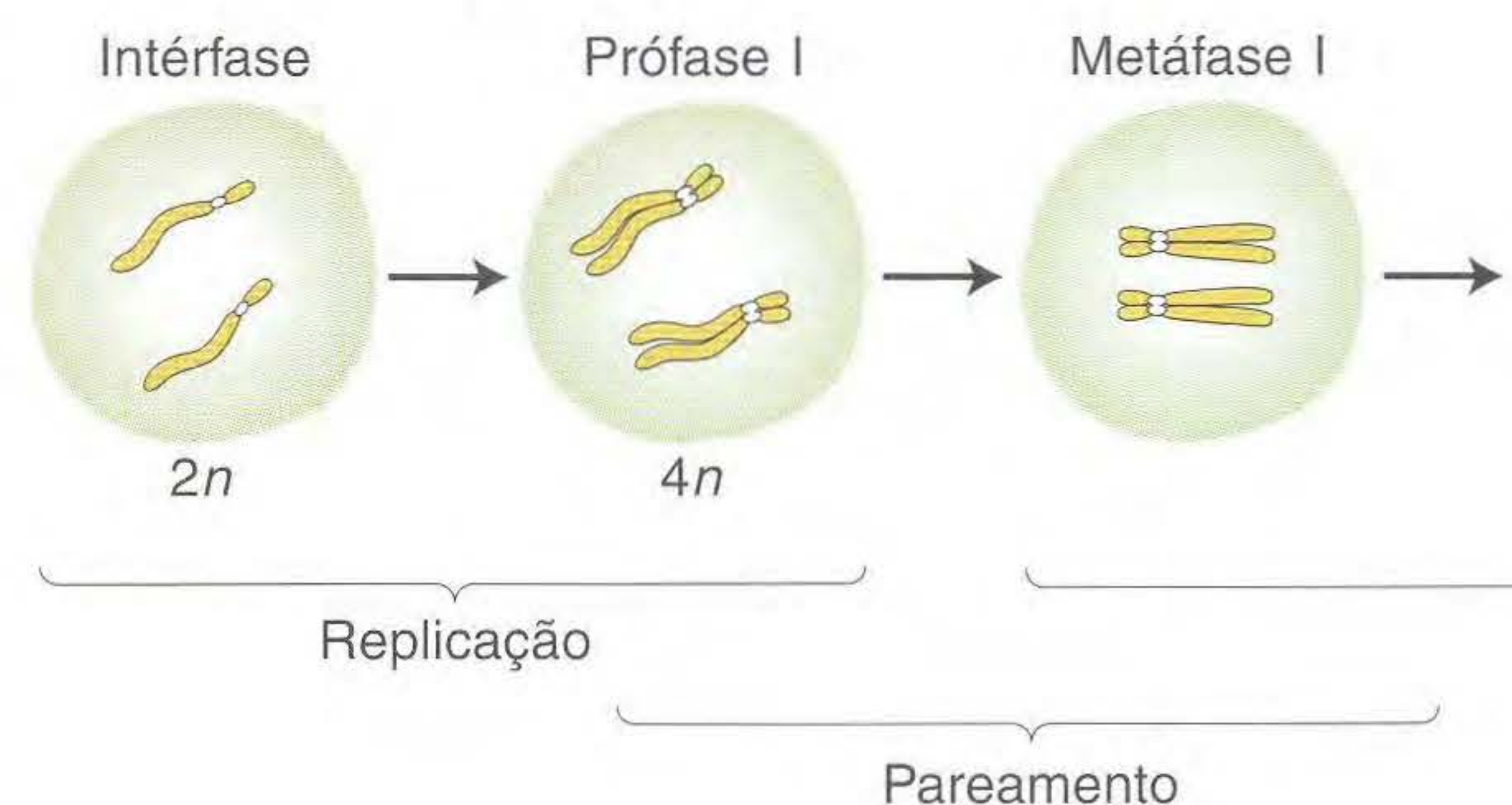
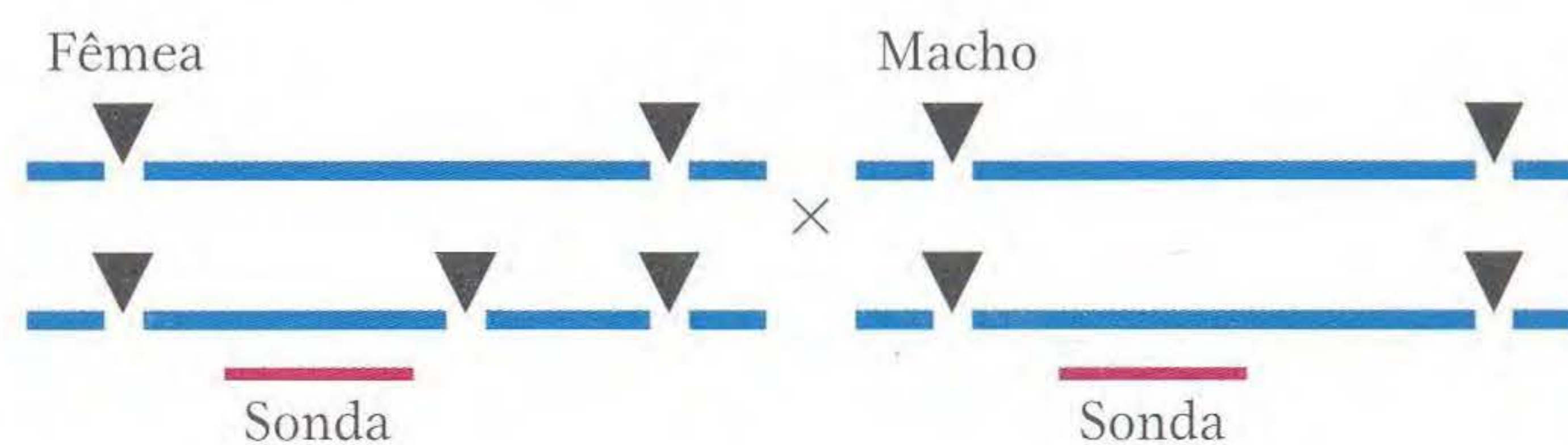
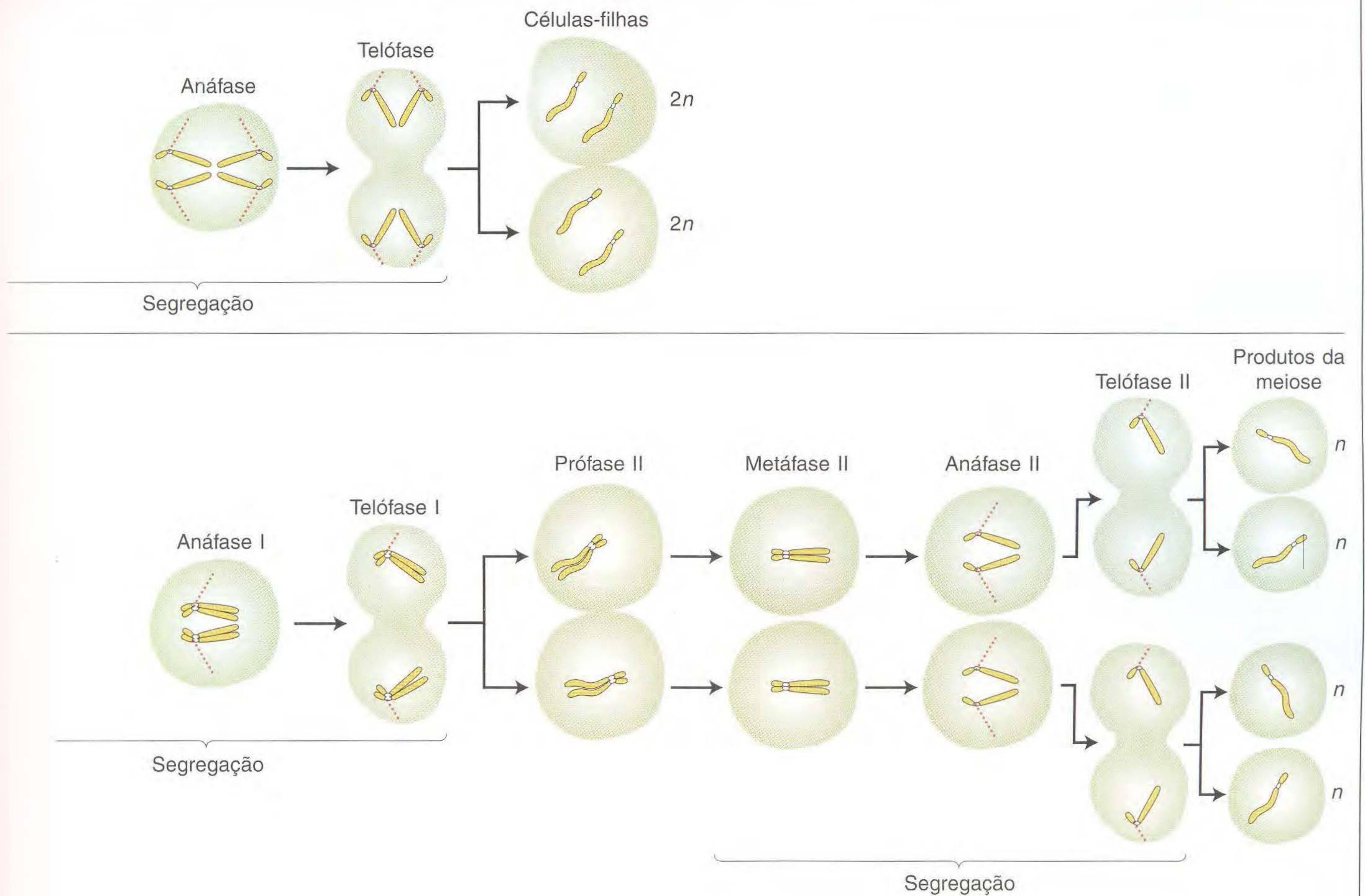


FIG. 2.15 Representação simplificada da mitose e meiose nas células diplóides ($2n$, diplóide; n , haplóide). (Versões detalhadas são mostradas no Boxe 2.1, mais adiante.)

cias-alvo não têm significado biológico em outros organismos que não bactérias. Eles estão presentes puramente ao acaso. Embora os sítios-alvo sejam geralmente encontrados consistentemente em locais específicos, às vezes, em qualquer cromossomo, um sítio-alvo específico está faltando ou existe um sítio extra. Se tal sítio flanqueia a seqüência hibridizada com a sonda, então uma hibridização de Southern revelará um RFLP. Considere este exemplo simples no qual um cromossomo de um genitor contém um sítio extra não encontrado nos outros cromossomos desse tipo nesse cruzamento:



As hibridizações de Southern mostrarão duas bandas na fêmea e apenas uma no macho. Os fragmentos “heterozí-



Complexos sinaptonêmicos na meiose

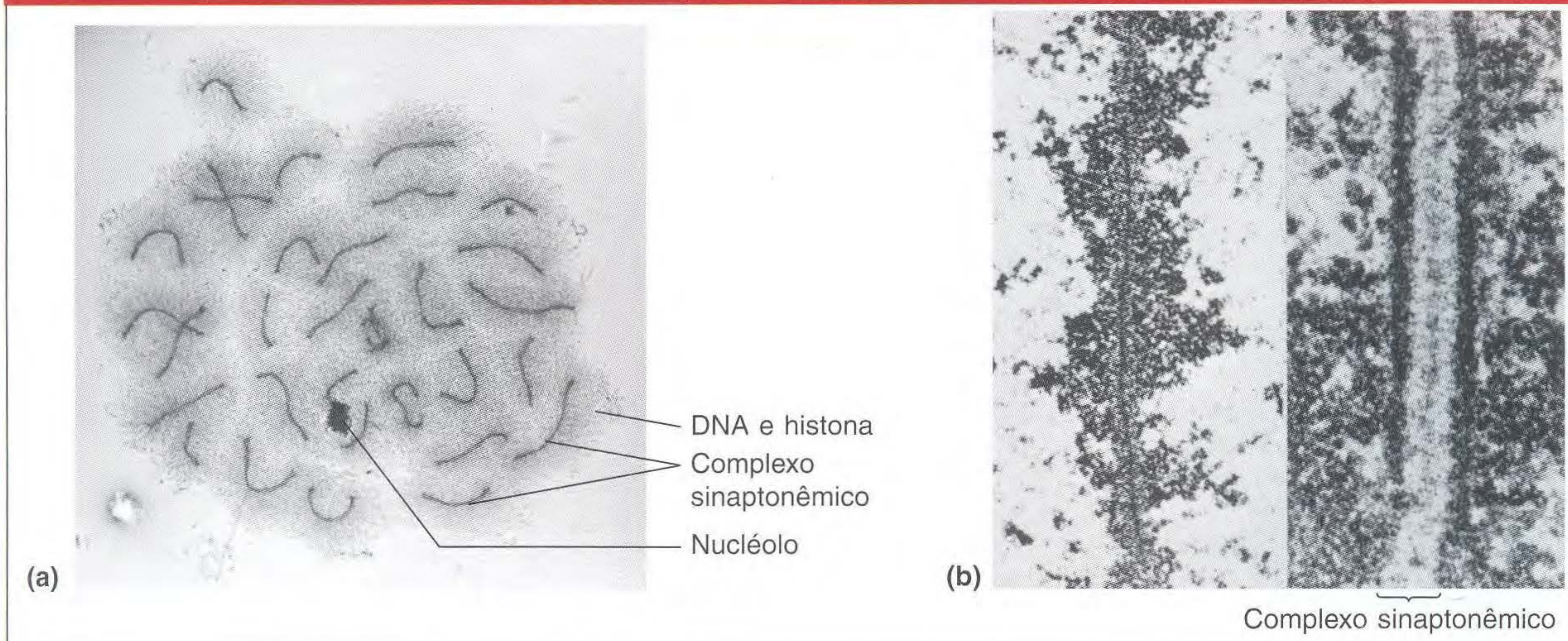


FIG. 2.16 (a) Em *Hyalophora cecropia*, um bicho-da-seda, o número cromossômico masculino normal é 62, dando 31 complexos sinaptonêmicos. No indivíduo mostrado aqui, um cromossomo (*centro*) é representado três vezes; tal cromossomo é chamado de *trivalente*. O DNA é disposto em alças regulares ao redor do complexo sinaptonêmico. (b) Complexo sinaptonêmico regular em *Lilium tyrinum*. Note (*direita*) os dois elementos laterais do complexo sinaptonêmico e (*esquerda*) um cromossomo não-pareado, mostrando um cerne correspondente a um dos elementos laterais. [Cortesia de Peter Moens.]

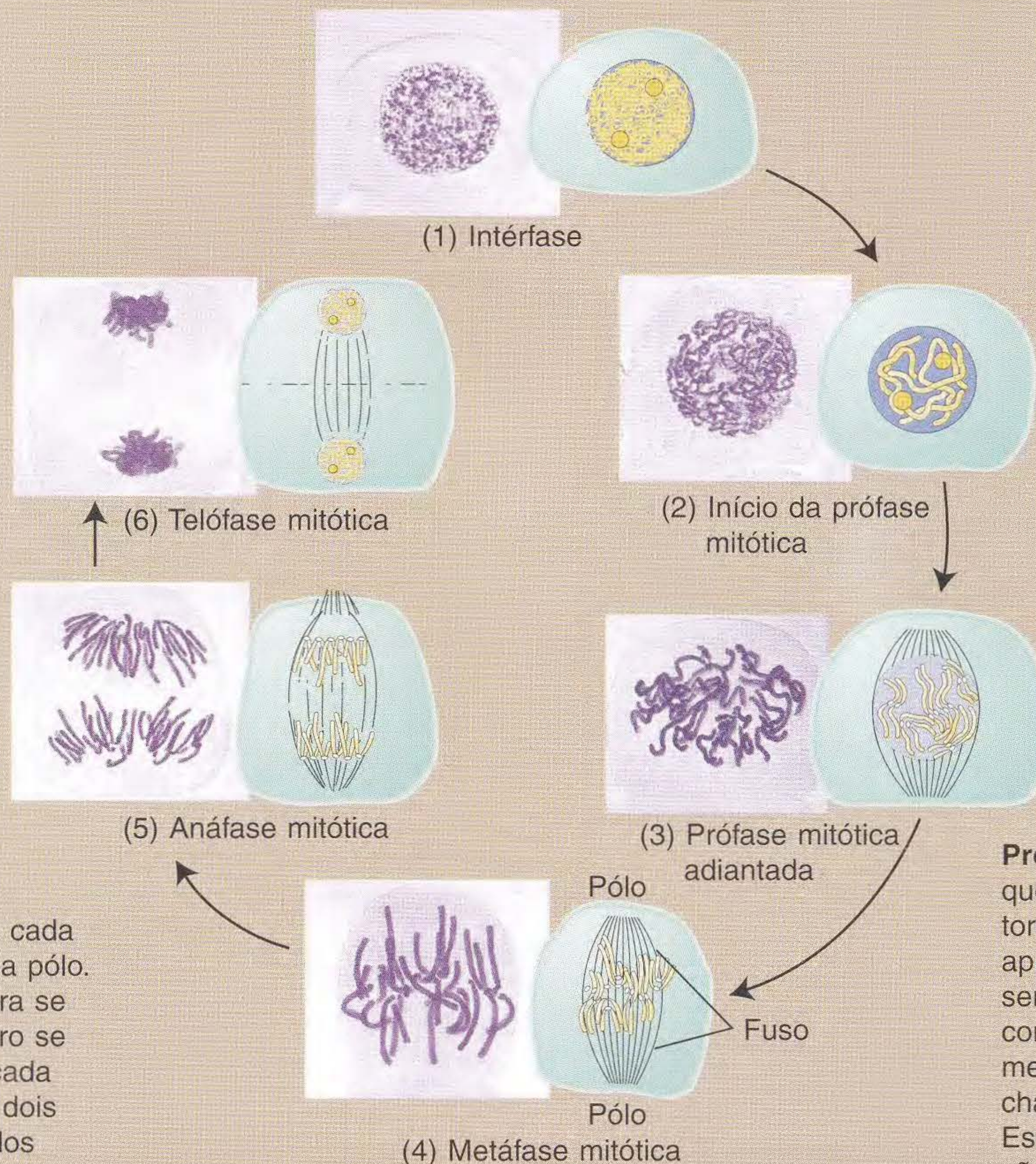
Boxe 2.1**Estágios da Mitose**

A mitose geralmente leva apenas uma pequena proporção do ciclo celular, aproximadamente 5 a 10%. O tempo restante é da intérfase, composta dos estágios G1, S e G2. O DNA é replicado durante a fase S, embora o DNA duplicado não se torne visível até a mitose adiantada. Os cromossomos não podem ser vistos durante a intérfase

(veja foto), principalmente porque estão em um estado distendido e entrelaçados.

As fotos a seguir mostram os estágios da mitose nos núcleos das células da ponta da raiz de lírio real, *Lilium regale*. Em cada estágio é mostrada uma foto com sua interpretação.

Telófase. A membrana nuclear se reestrutura ao redor de cada núcleo-filho, os cromossomos se deslicoidizam e os nucléolos reaparecem, havendo a reconstituição dos núcleos interfásicos. Ao final da telófase, o fuso se dispersou, e o citoplasma foi dividido em dois por uma nova membrana celular.

**Começo da prófase.**

Os cromossomos tornam-se distintos pela primeira vez. Eles ficam progressivamente mais curtos por um processo de contração, ou condensação, em uma série de hélices; as hélices produzem estruturas que são mais facilmente movimentadas.

Prófase adiantada. À medida que os cromossomos se tornam visíveis, eles se apresentam bifilamentares, sendo cada cromossomo composto de duas metades longitudinais chamadas cromátides. Essas cromátides-“irmãs” são unidas no centrômero. Os nucléolos, estruturas esféricas intranucleares grandes, desaparecem nesse estágio. A membrana nuclear começa a se desfazer, e o nucleoplasma e o citoplasma tornam-se um.

Anáfase. Os pares de cromátides-irmãs separam-se, com uma de cada par movendo-se para cada pólo. Os centrômeros, que agora se mostram divididos, primeiro se separam. À medida que cada cromátide se move, seus dois braços parecem ser levados pelo centrômero; resulta uma série de estruturas em forma de V, com as pontas dos V dirigidas para os pólos.

Metáfase. O fuso nuclear torna-se proeminente. O fuso é uma estrutura em rede que se forma na área nuclear; ele consiste em uma série de fibras paralelas que se dirigem para cada um dos dois pólos celulares. Os cromossomos movem-se para o plano equatorial da célula, onde os centrômeros se ligam às fibras do fuso de cada pólo.

As fotos mostram a mitose nos núcleos das células da ponta da raiz de *Lilium regale*.

[De J. McLeish and B. Snode, *Looking at Chromosomes*. Copyright 1958, St. Martin's, Macmillan.]

gotos” serão herdados exatamente do mesmo modo que um gene. Os resultados do cruzamento anterior poderiam ser escritos do seguinte modo:

longo/curto \times longo/longo

e a prole seria

$\frac{1}{2}$ longo/curto

$\frac{1}{2}$ longo/longo

de acordo com a lei da segregação.

No exemplo precedente, o RFLP não está associado a nenhuma diferença fenotipicamente mensurável, e os sítios

Demonstração de segregação igual dentro de um meiócito na levedura *S. cerevisiae*

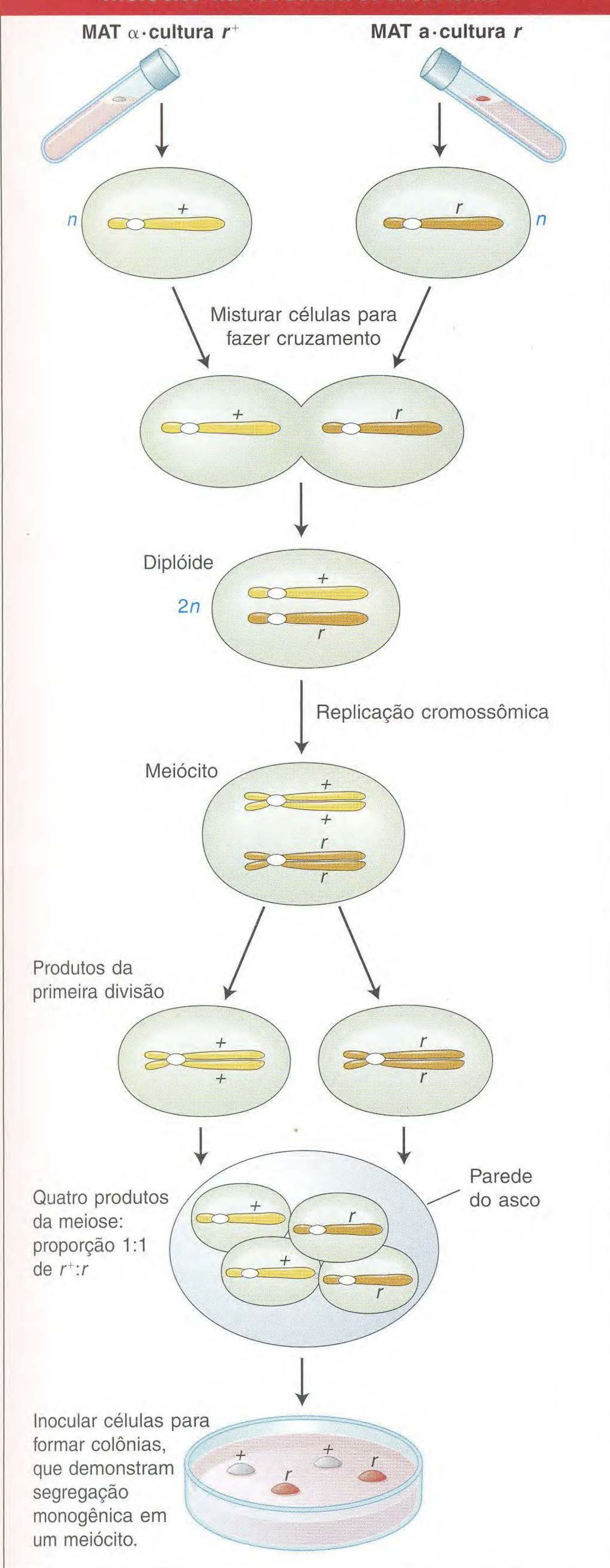
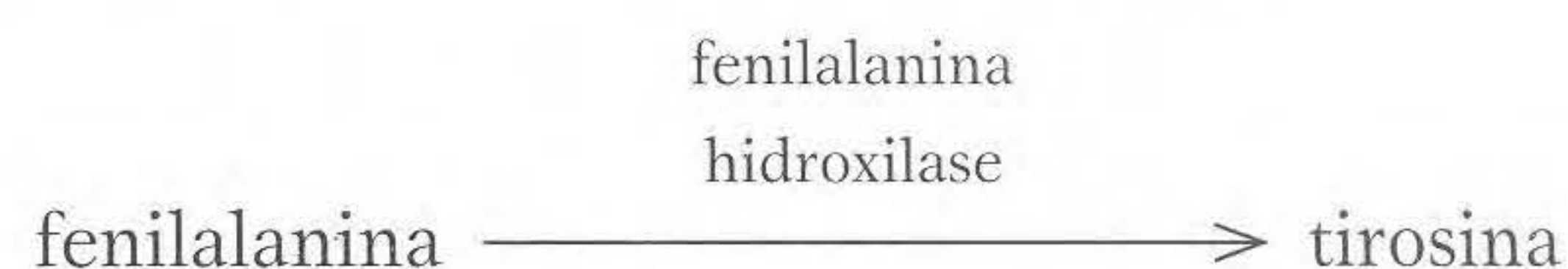


FIG. 2.17 Um asco isolado do cruzamento $+ \times r$ leva a duas culturas de $+$ e duas de r .

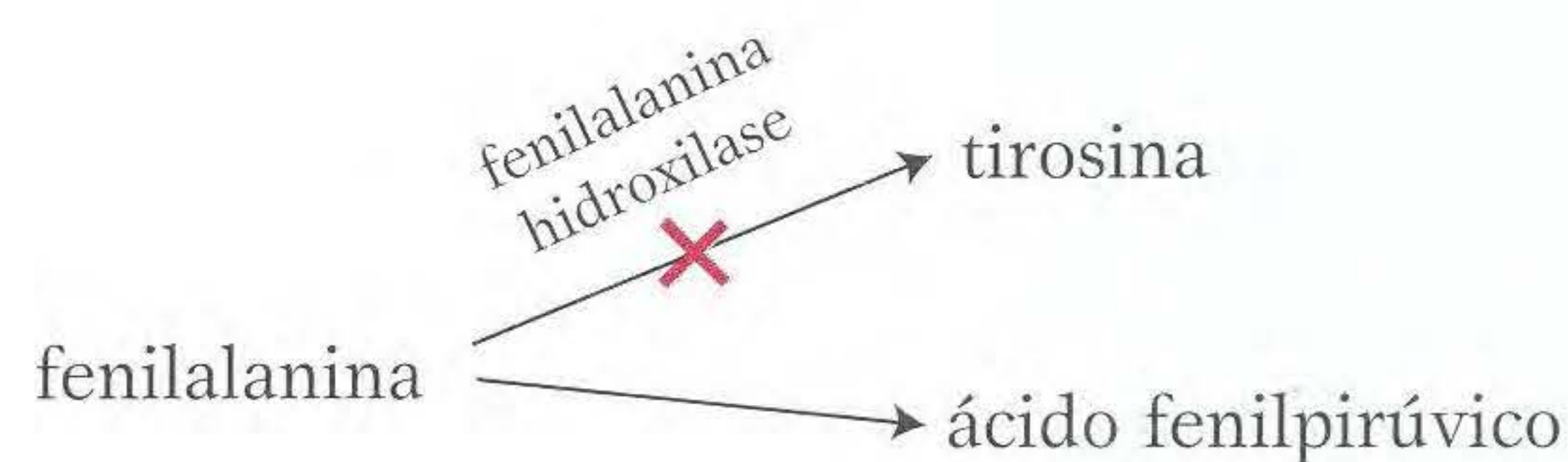
podem estar dentro de um gene ou não. Entretanto, às vezes, por acaso, uma mutação dentro da sequência codificante de um gene que produz um fenótipo mutante também introduz um novo sítio-alvo para uma enzima de restrição. O novo sítio-alvo fornece uma marcação molecular conveniente para os alelos. Por exemplo, uma mutação recessiva em um gene para a síntese de pigmento pode produzir um alelo mutante albino e, ao mesmo tempo, um novo sítio de restrição. Assim, uma sonda para o alelo mutante detecta dois fragmentos em a e apenas um em A . O padrão de herança é mostrado na Figura 2.20. Nessa ilustração, vemos uma demonstração direta de segregação molecular ocorrendo junto com segregação alélica e fenotípica. (Outros tipos de mutações produziram efeitos diferentes ao nível detectado pelas análises de Southern, Northern e Western.)

A natureza dos alelos e seus produtos Usamos o conceito de *alelos* sem defini-los a nível molecular. Quais são as *diferenças estruturais* entre o tipo selvagem e os alelos mutantes ao nível do DNA de um gene? Quais são as *diferenças funcionais* ao nível de proteína? Os alelos mutantes podem ser usados para estudar a herança monogênica sem a necessidade de compreender sua natureza estrutural ou funcional. Entretanto, como um motivo primário para usar uma herança monogênica é ultimamente investigar a função de um gene, devemos analisar a natureza molecular dos alelos tipo selvagem e mutante tanto em níveis estruturais quanto funcionais.

Vamos explorar o tópico usando a doença humana fenilcetonúria (PKU). Veremos, em uma seção posterior sobre análise de heredogramas, que o fenótipo PKU é herdado como recessivo mendeliano. A doença é causada por um alelo defeituoso do gene que codifica a enzima hepática fenilalanina hidroxilase (PAH). Essa enzima normalmente converte a fenilalanina em alimento no aminoácido tirosina:



Entretanto, uma mutação no gene que codifica essa enzima pode alterar a sequência de aminoácidos na vizinhança do sítio ativo da enzima. Nesse caso, a enzima não pode ligar-se à fenilalanina (seu substrato) nem convertê-la em tirosina. Portanto, a fenilalanina acumula-se no corpo e é convertida em ácido fenilpirúvico. Esse composto interfere no desenvolvimento do sistema nervoso, levando a retardo mental.

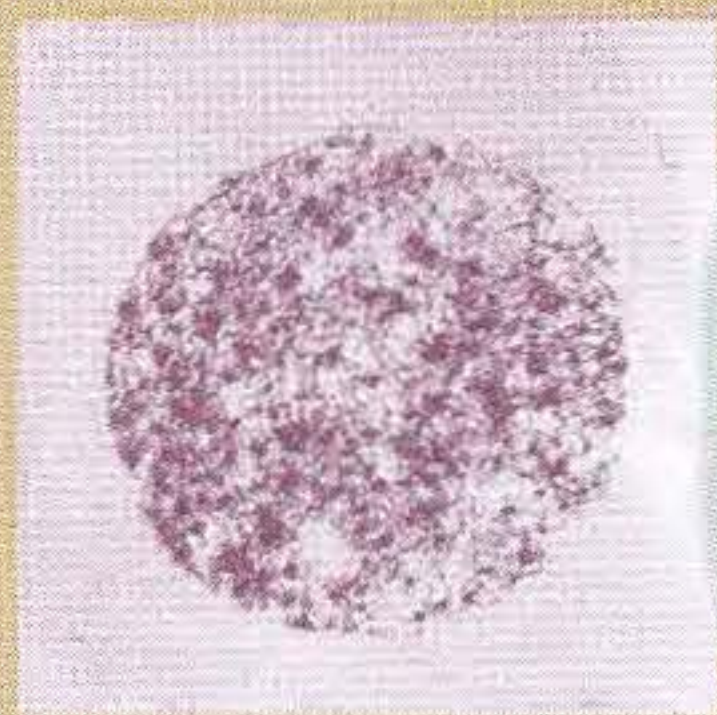


As crianças hoje são rotineiramente testadas quanto a essa deficiência de processamento ao nascimento. Se a deficiência é detectada, a fenilalanina pode ser evitada com o uso de uma dieta especial, e o desenvolvimento da doença será detido.

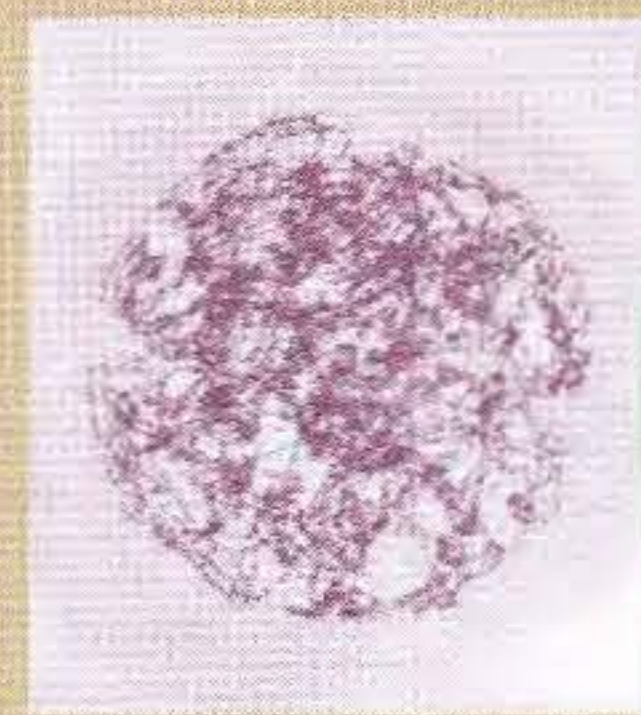
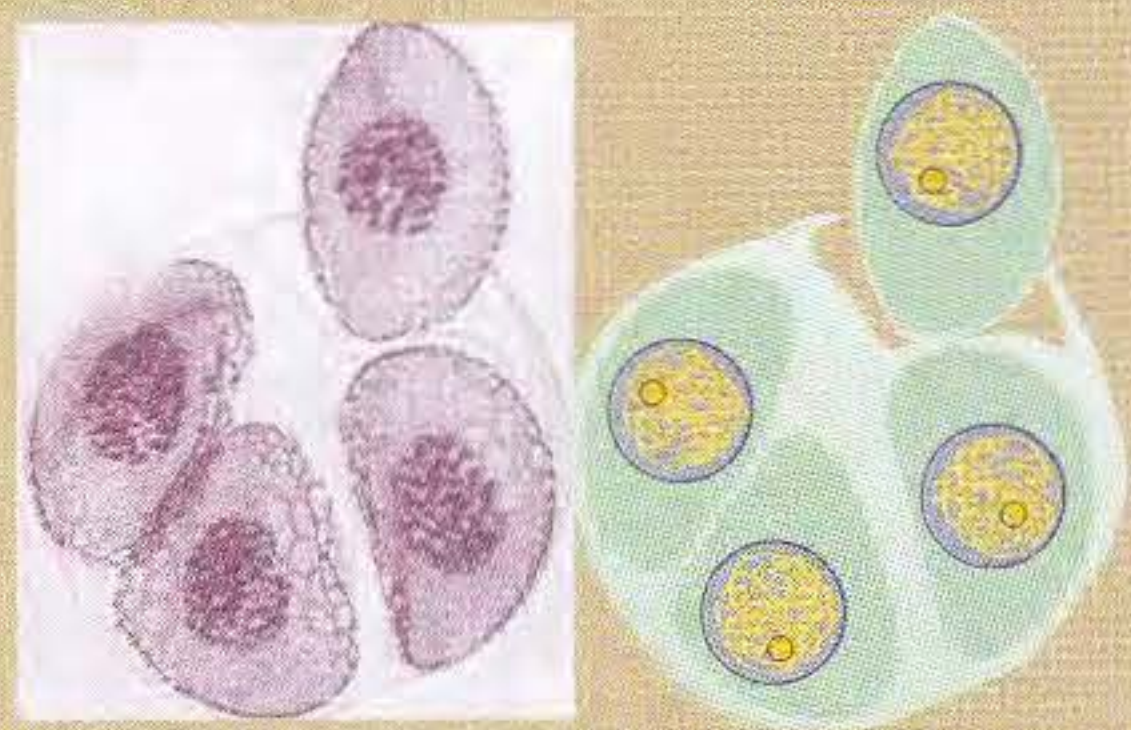
Boxe 2.2**Estágios da Meiose**

A meiose consiste em duas divisões nucleares distintas como meiose I e meiose II, que ocorrem em divisões celulares consecutivas. Cada divisão meiótica é dividida em prófase, metáfase, anáfase e telófase. Desses estágios, o mais complexo e longo é a prófase I, que é dividida em cinco estágios.

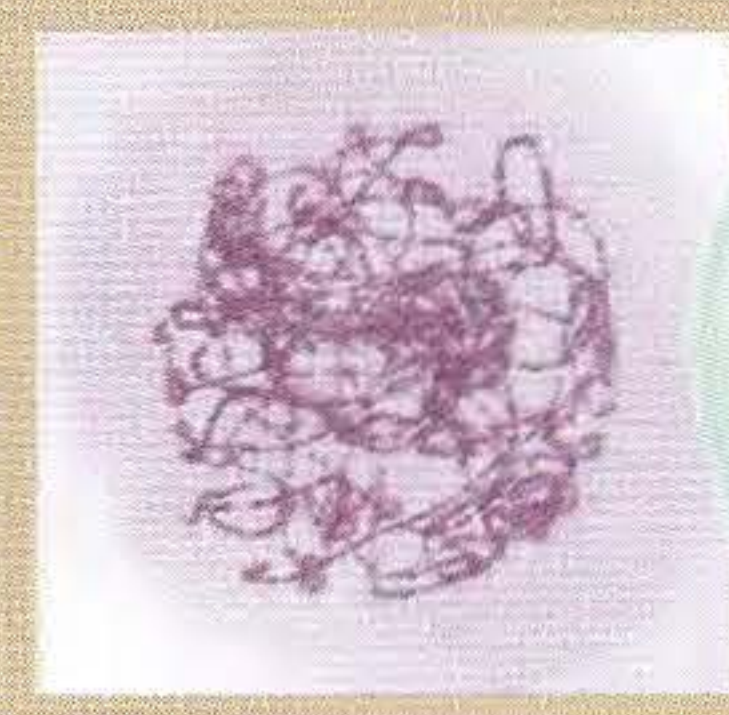
As fotos a seguir mostram os estágios da meiose nos núcleos das células da ponta da raiz de lírio real, *Lilium regale*. Em cada estágio é mostrada uma foto e um desenho com sua interpretação.

**1 Leptóteno**

Prófase I: Leptóteno. Os cromossomos tornam-se visíveis como finos filamentos longos. Os cromossomos continuam a se contrair no leptóteno e durante toda a prófase. Pequenas áreas de espessamento (cromômeros) desenvolvem-se ao longo de cada cromossomo, dando-lhe o aspecto de um colar de contas.

**2 Zigóteno**

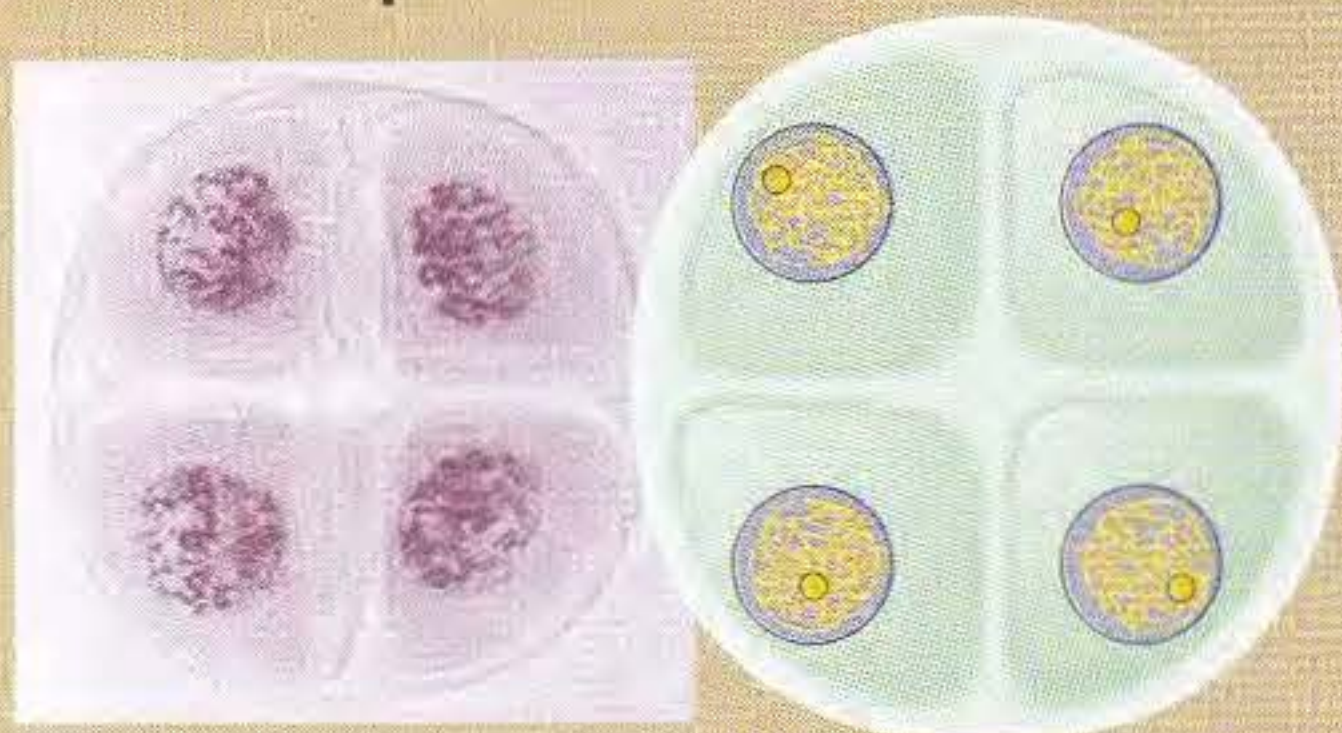
Prófase I: Zigóteno. O pareamento ativo dos filamentos torna aparente que o complemento cromossômico do meiócito é de fato dois conjuntos cromossômicos completos. Assim, cada cromossomo tem um parceiro de pareamento, e os dois tornam-se progressivamente pareados, ou em sinapse, ao longo de sua extensão.

**3 Paquíteno**

Prófase I: Paquíteno. Os cromossomos são espessos e em sinapse. Assim, o número de pares de cromossomos homólogos no núcleo é igual ao número n . Os nucléolos são, em geral, pronunciados durante o paquíteno. Os cromossomos tipo colar de contas alinham-se precisamente nos homólogos pareados, produzindo um padrão distinto para cada par.

**16 Jovens grãos de pólen**

A tétrade e os jovens grãos de pólen: Nas anteras de uma flor, os quatro produtos da meiose desenvolvem-se em grãos de pólen. Em outros organismos, a diferenciação produz outros tipos de estruturas a partir dos produtos de meiose, tais como espermatozoides em animais.

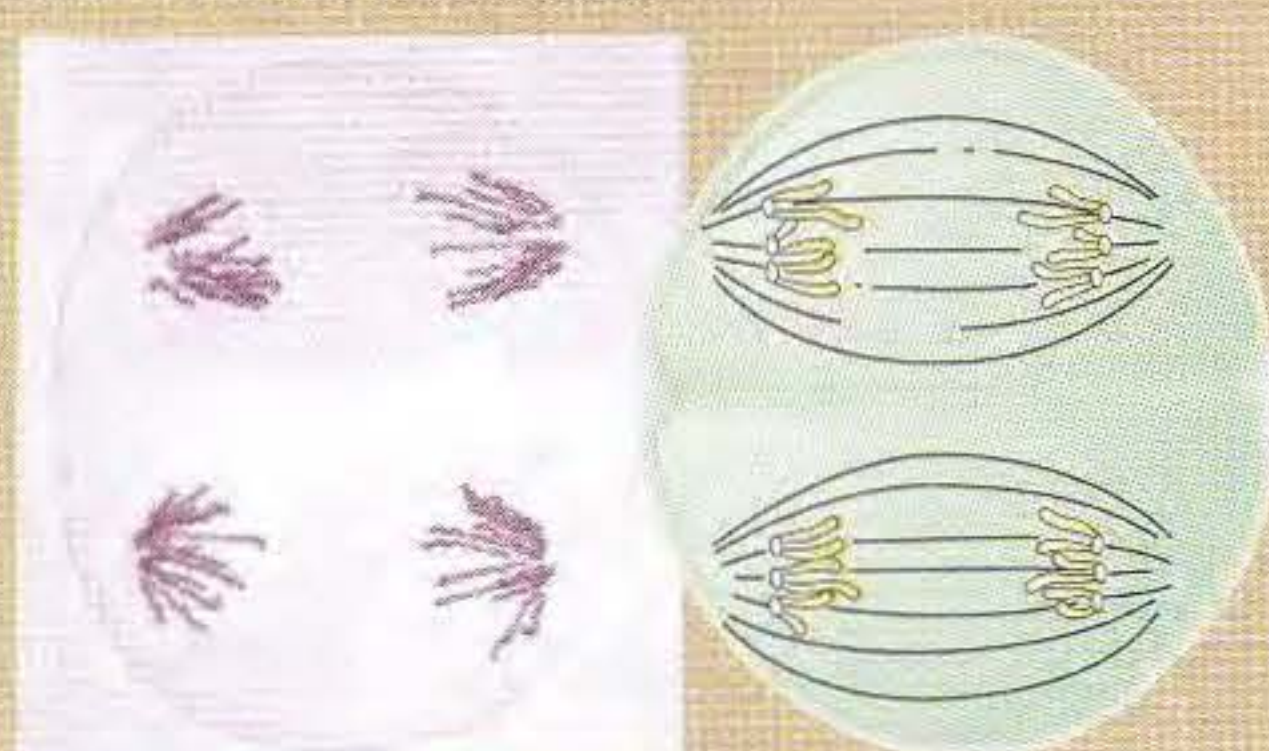
**15 A tétrade**

As células dividem-se

**14 Telófase II**

Telófase II: Os núcleos reconstituem-se ao redor dos cromossomos nos pólos.

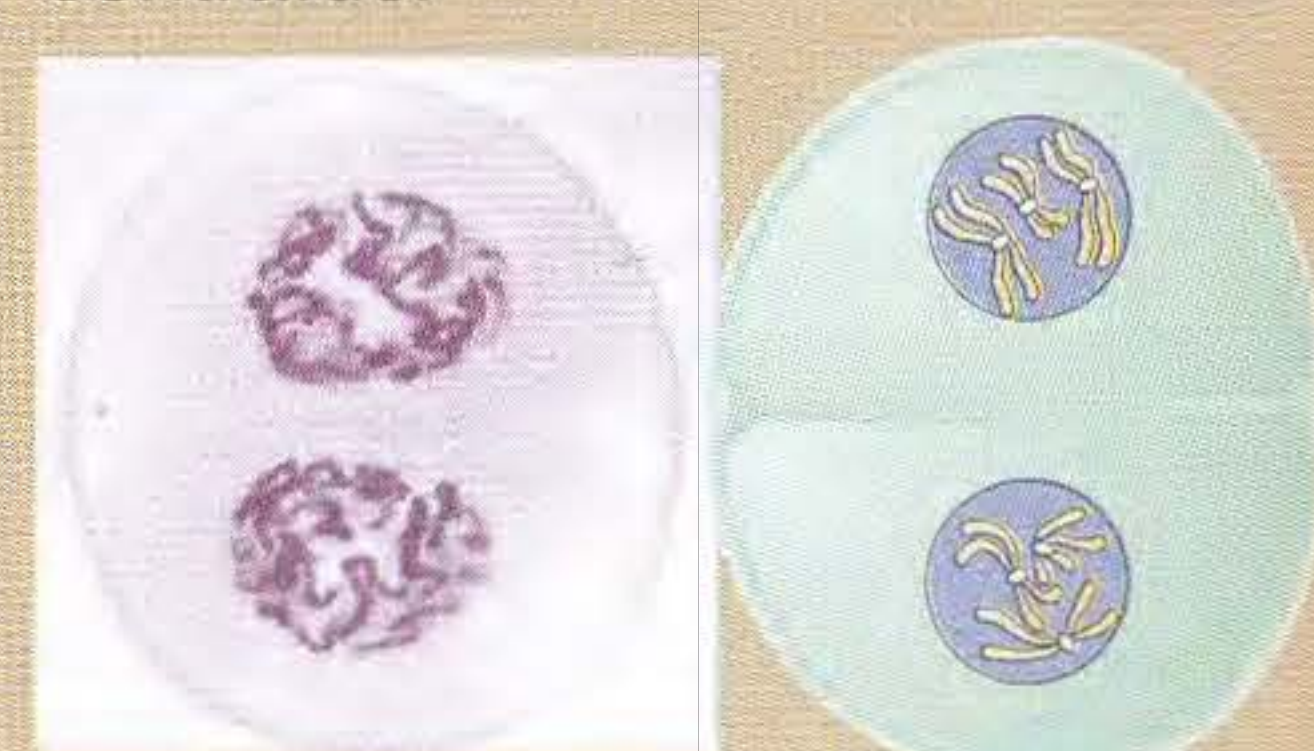
Anáfase II: Os centrômeros dividem-se, e as cromátides-irmãs são levadas para pólos opostos pelas fibras do fuso.

**13 Anáfase II**

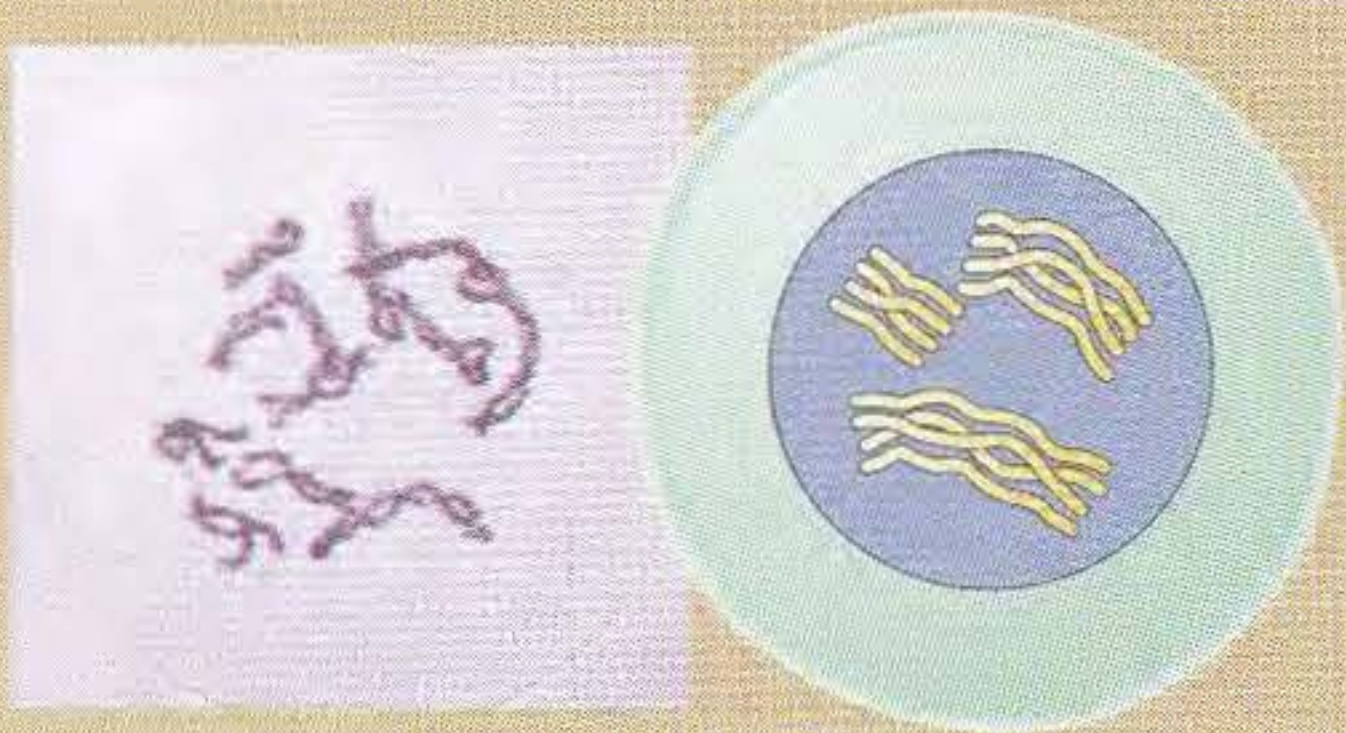
Metáfase II: Os pares de cromátides-irmãs arrumam-se no plano equatorial. Aqui, as cromátides em geral se dissociam parcialmente uma da outra, ao invés de ficarem bem juntas como na mitose.

**12 Metáfase II**

Prófase II: O número haplóide de pares de cromátides-irmãs agora está presente no estado contraído.

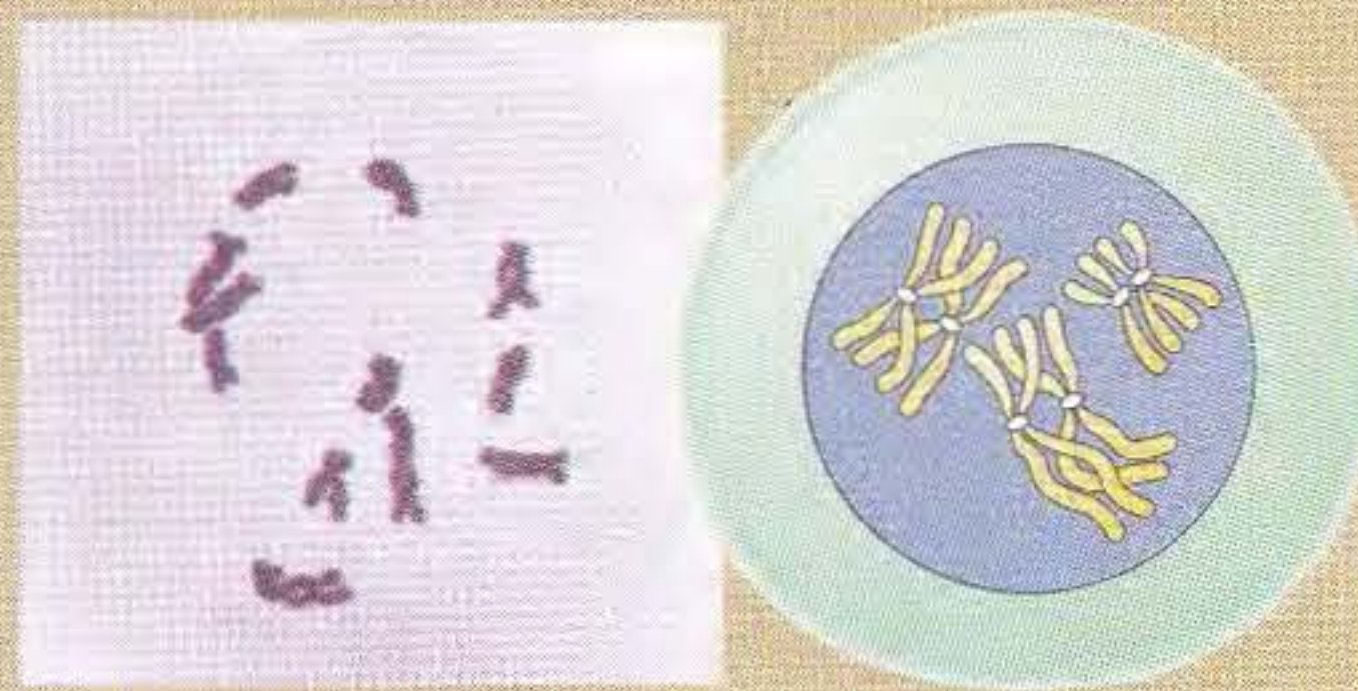
**11 Prófase II**

As fotos mostram meiose e a formação de pólen em *Lilium regale*. Note: Para simplificação, os vários quiasmas são desenhados entre apenas duas cromátides; de fato, todas as quatro cromátides podem participar. [De J. McLeish and B. Snode, *Looking at Chromosomes*. Copyright 1958, St. Martin's, Macmillan.]



4 Diplóteno

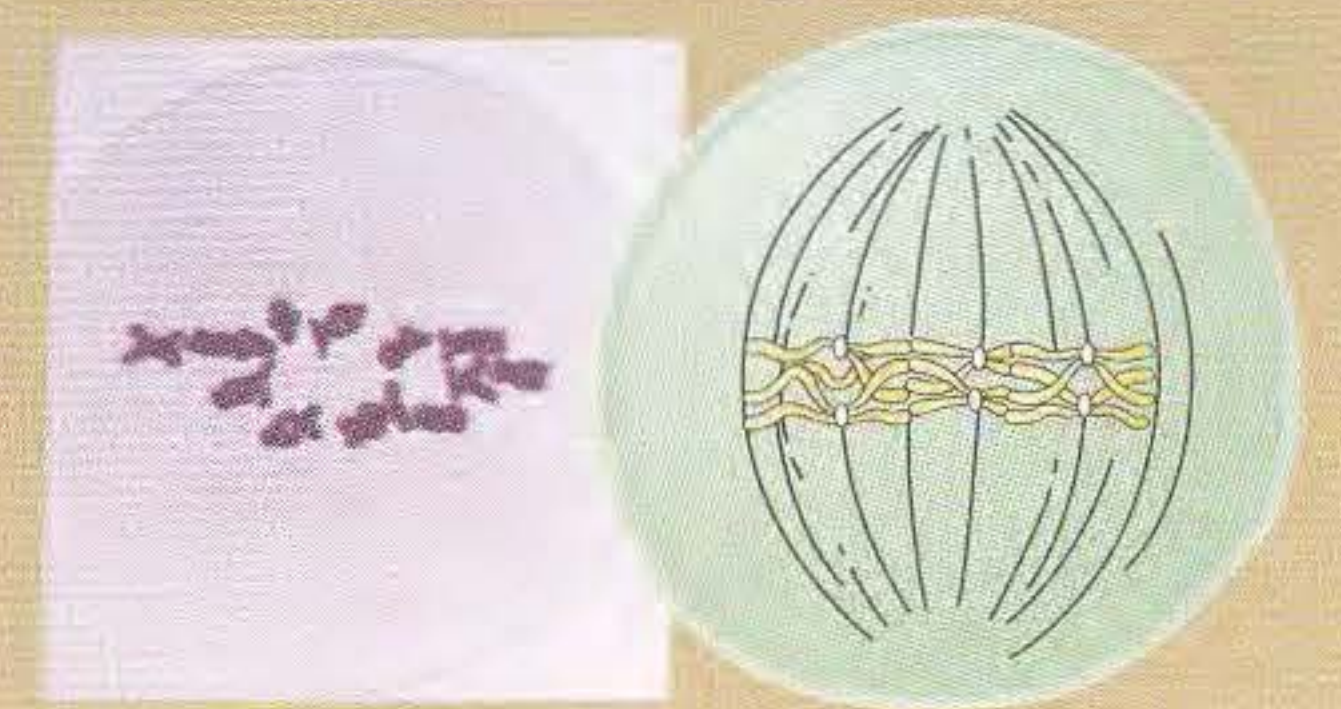
Prófase I: Diplóteno. Embora cada homólogo pareça ser um filamento único no leptóteno, o DNA já se replicou durante a fase S pré-meiótica. Esse fato manifesta-se no diplóteno à medida que cada cromossomo é visto tendo um par de cromátides-irmãs. A estrutura em sinapse agora consiste em um feixe de quatro cromátides homólogas. O pareamento entre os homólogos torna-se menos intenso; de fato, eles parecem repelir-se um ao outro, e, à medida que se separam um pouco, aparece uma estrutura em forma de cruz, chamada quiasma, entre as cromátides não-irmãs. Cada par de cromossomos em geral tem um ou mais quiasmas.



5 Diacinese

Prófase I: Diacinese. Esse estágio difere apenas um pouco do diplóteno, exceto pela maior contração cromossômica. Ao final da diacinese, o longo cromossomo filamentososo da interfase foi substituído por unidades compactas que são mais móveis.

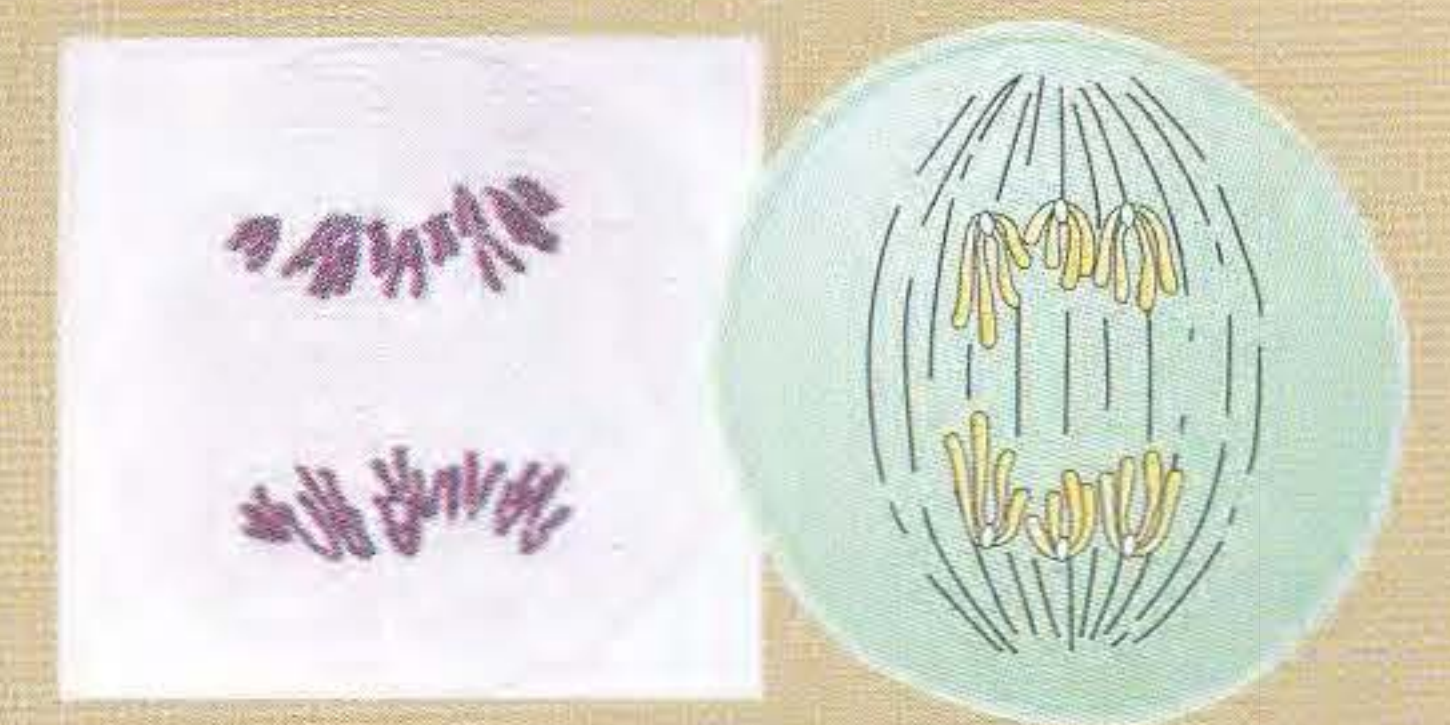
Metáfase I: A membrana nuclear e os nucléolos desapareceram, e cada par de homólogos fica em uma posição no plano equatorial. Nesse estágio da meiose, os centrômeros não se dividem. Essa falta de divisão é uma importante diferença da mitose. Os dois centrômeros de um par de cromossomos homólogos ligam-se às fibras do fuso de pólos opostos.



6 Metáfase I



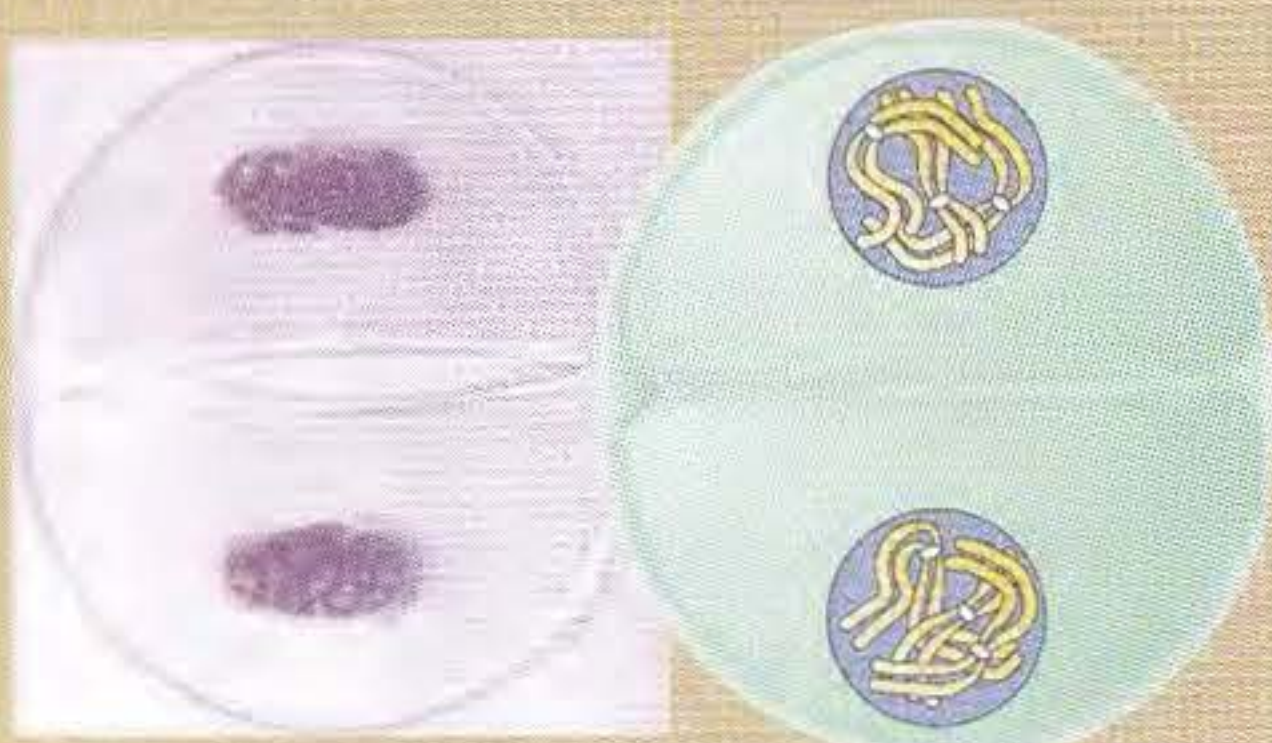
7 Início de anáfase I



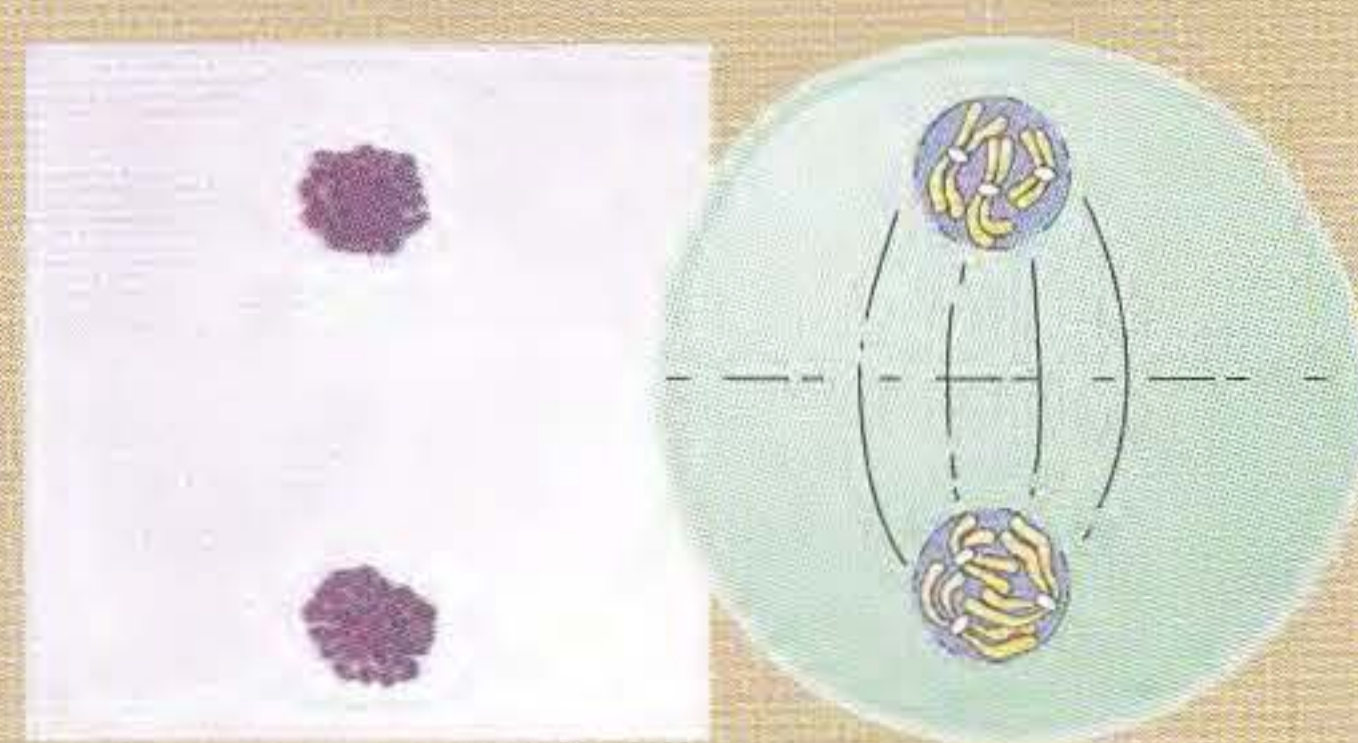
8 Anáfase I avançada

Anáfase I: A anáfase começa quando os cromossomos se movem em direção aos pólos. Os membros de um par de homólogos movem-se para pólos opostos.

Telófase I e interfase: Os cromossomos alongam-se e tornam-se difusos. A membrana nuclear se reconstitui, e a célula divide-se. Após a telófase I, há uma interfase, chamada intercinese. A telófase I e a intercinese não são universais. Em muitos organismos, esses estágios não existem, não há reconstituição de membrana nuclear e, após a divisão celular, as células vão diretamente para a meiose II. Em outros organismos, a telófase I e a intercinese têm duração curta. Em qualquer caso, não há síntese de DNA nessa época, e o estado genético dos cromossomos não muda.



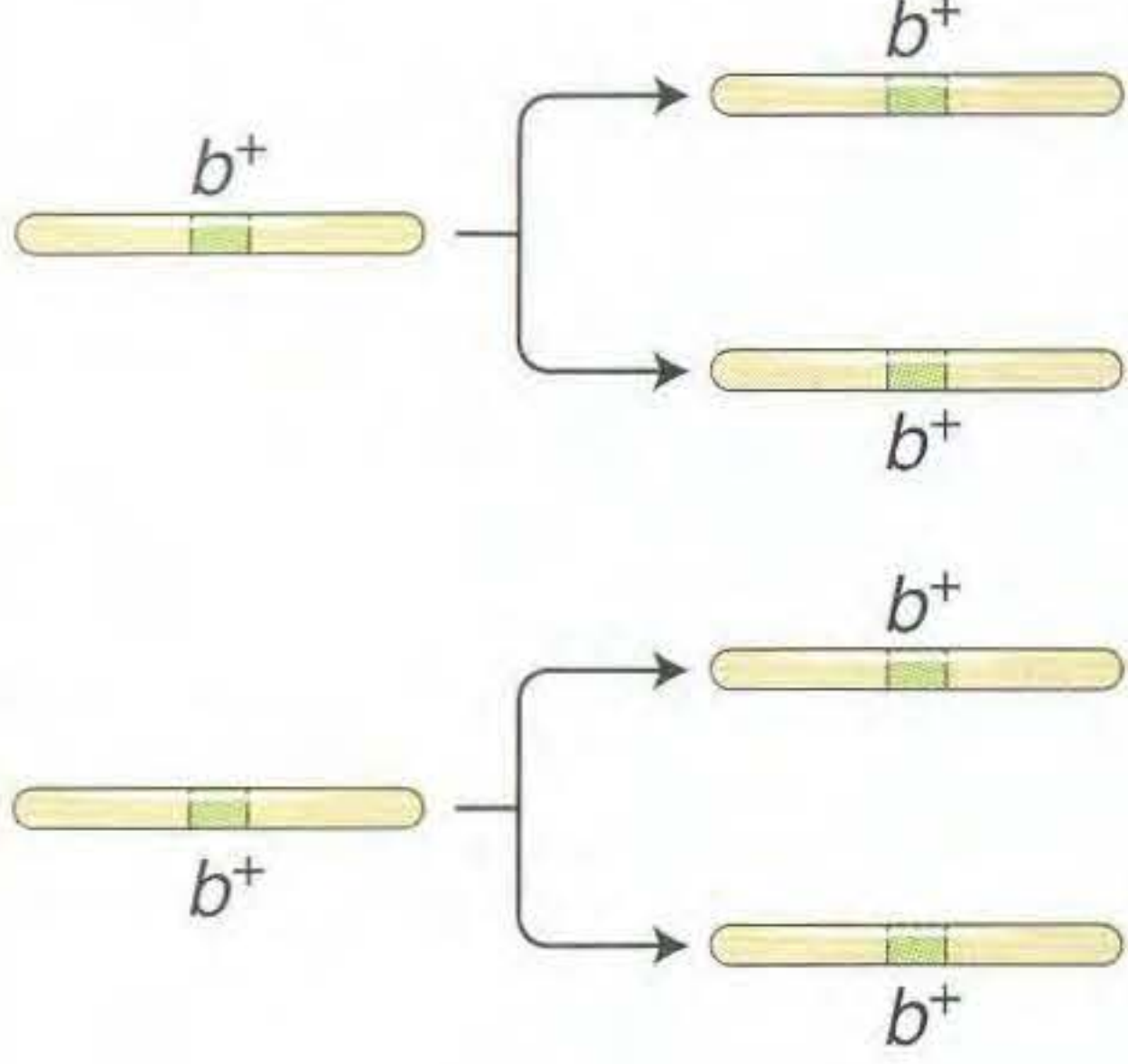
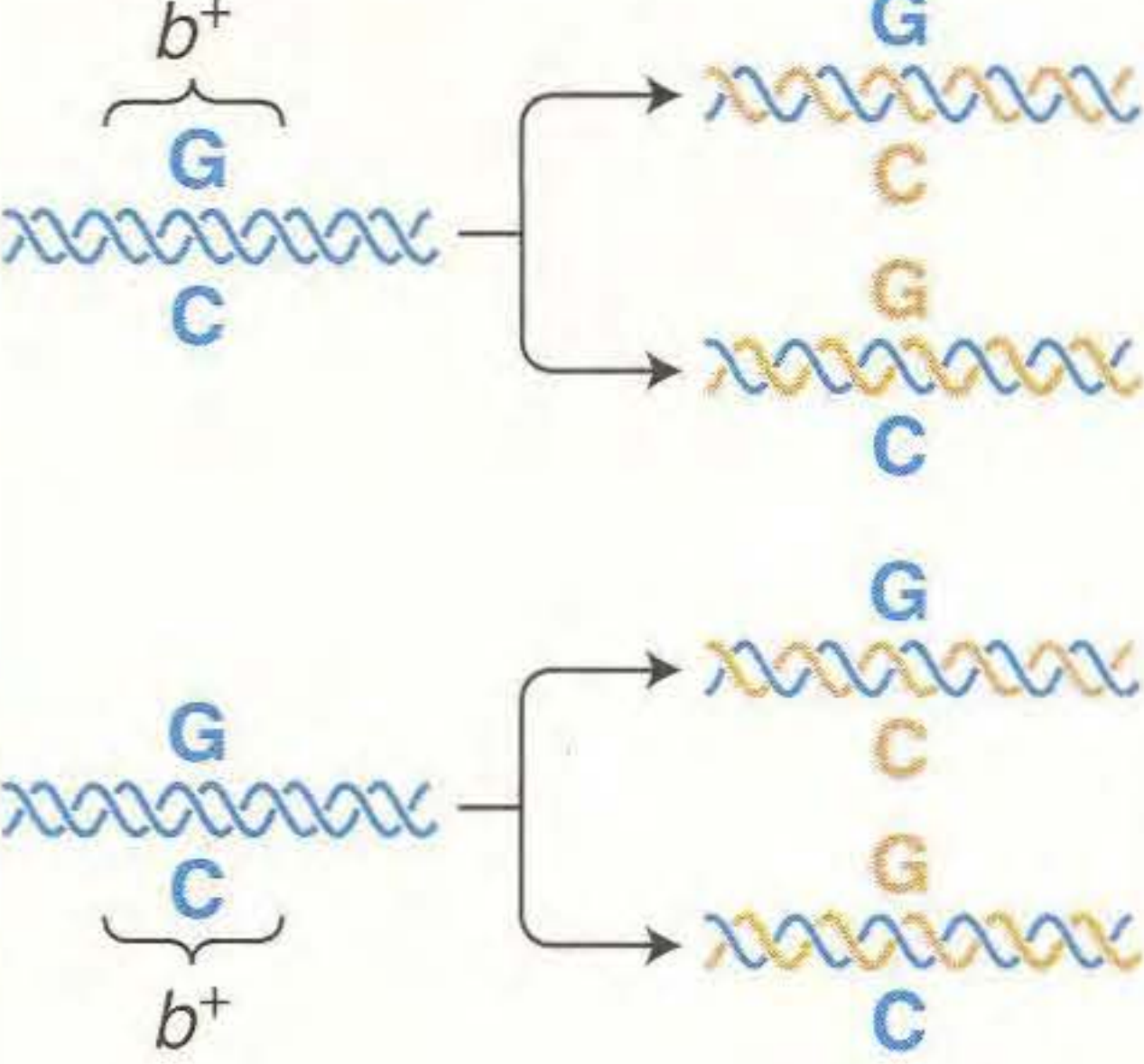
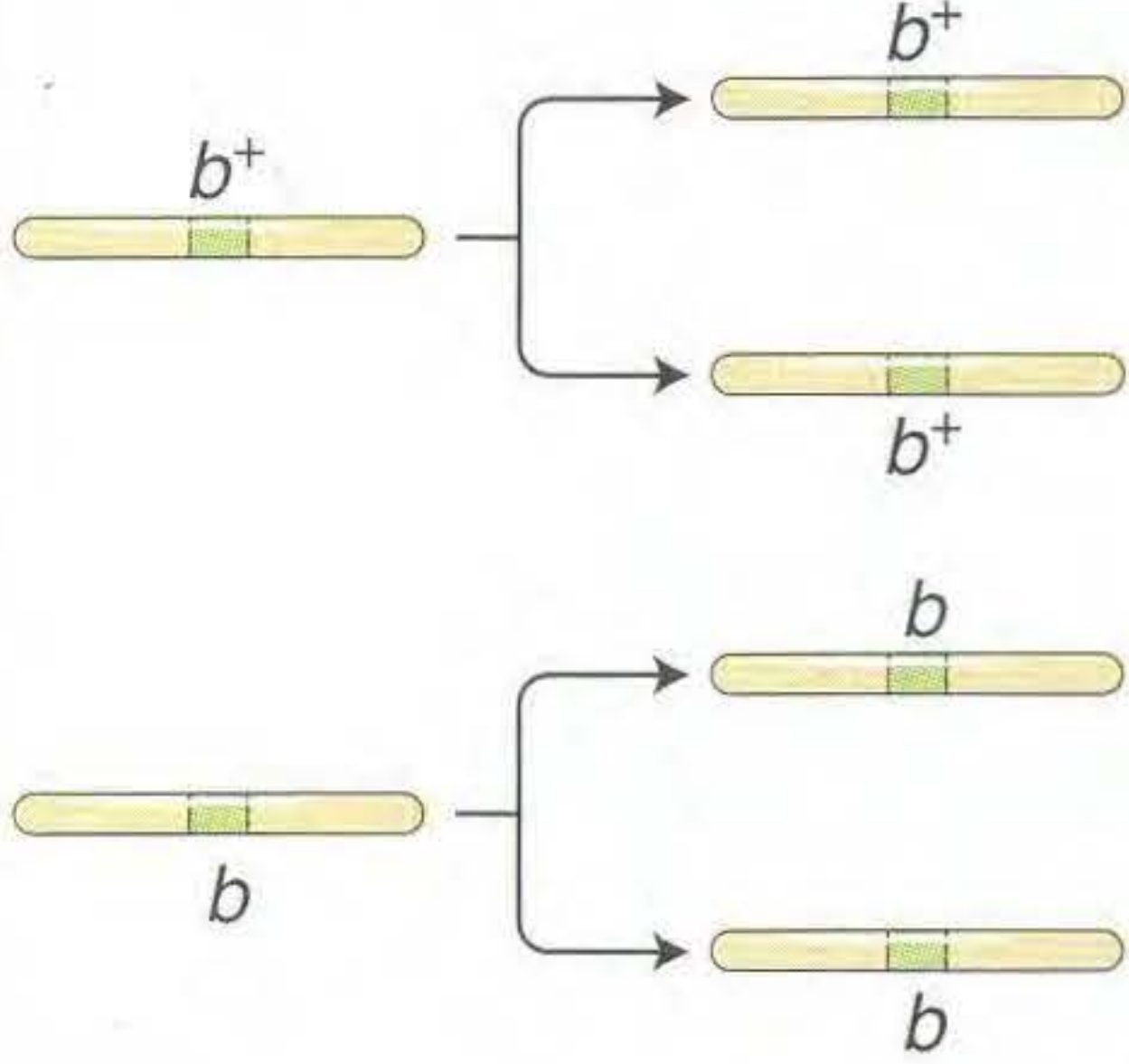
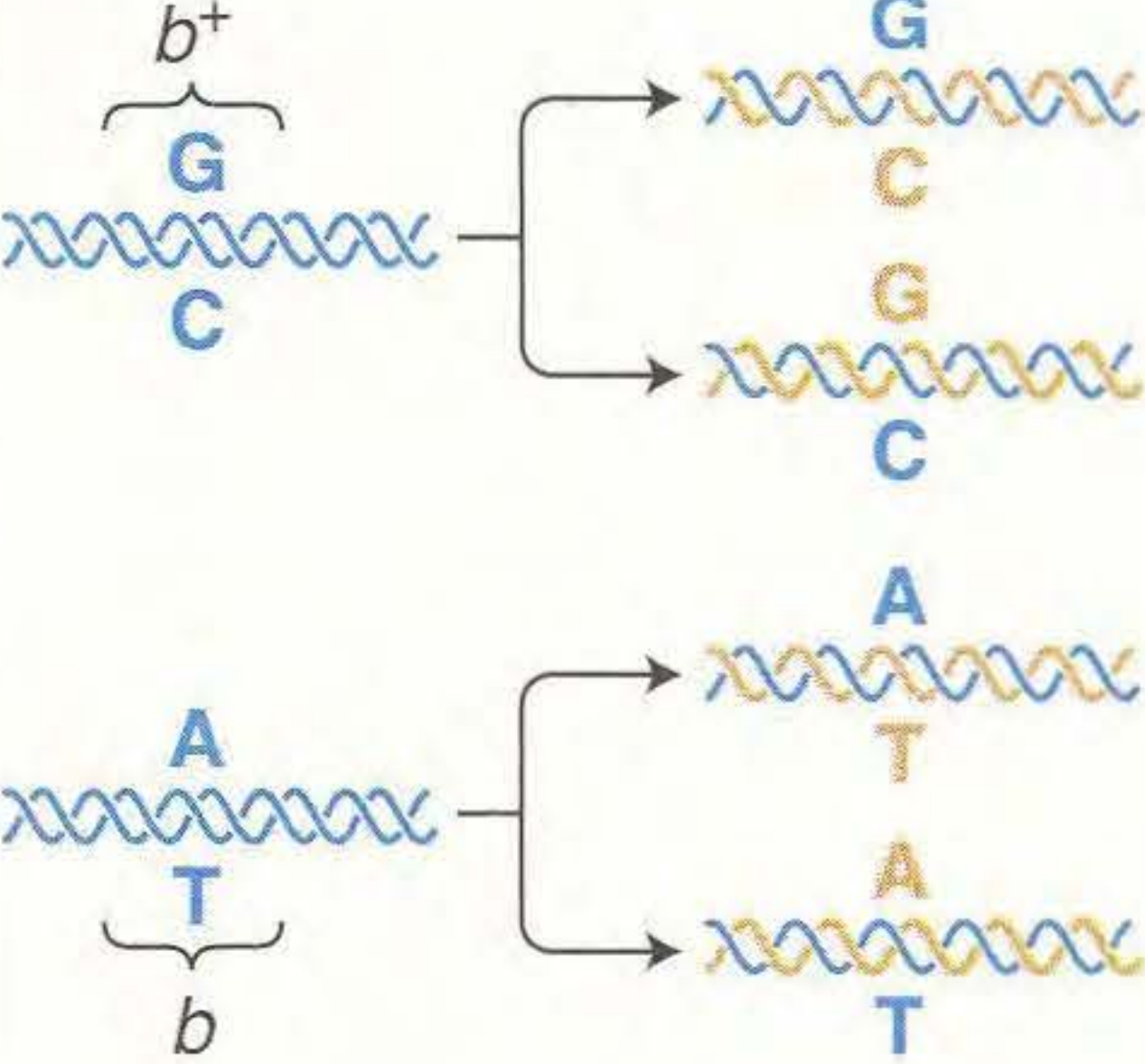
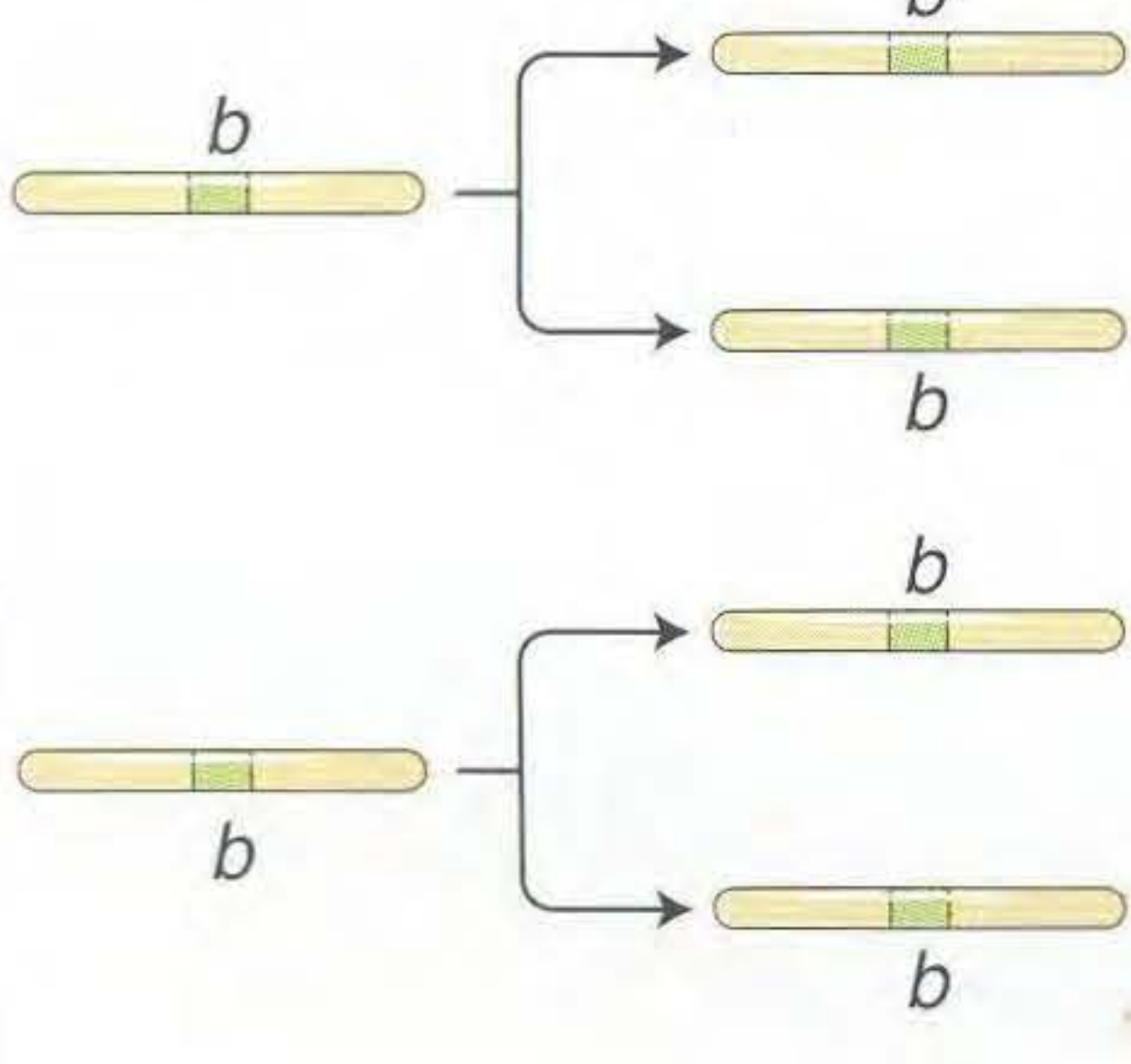
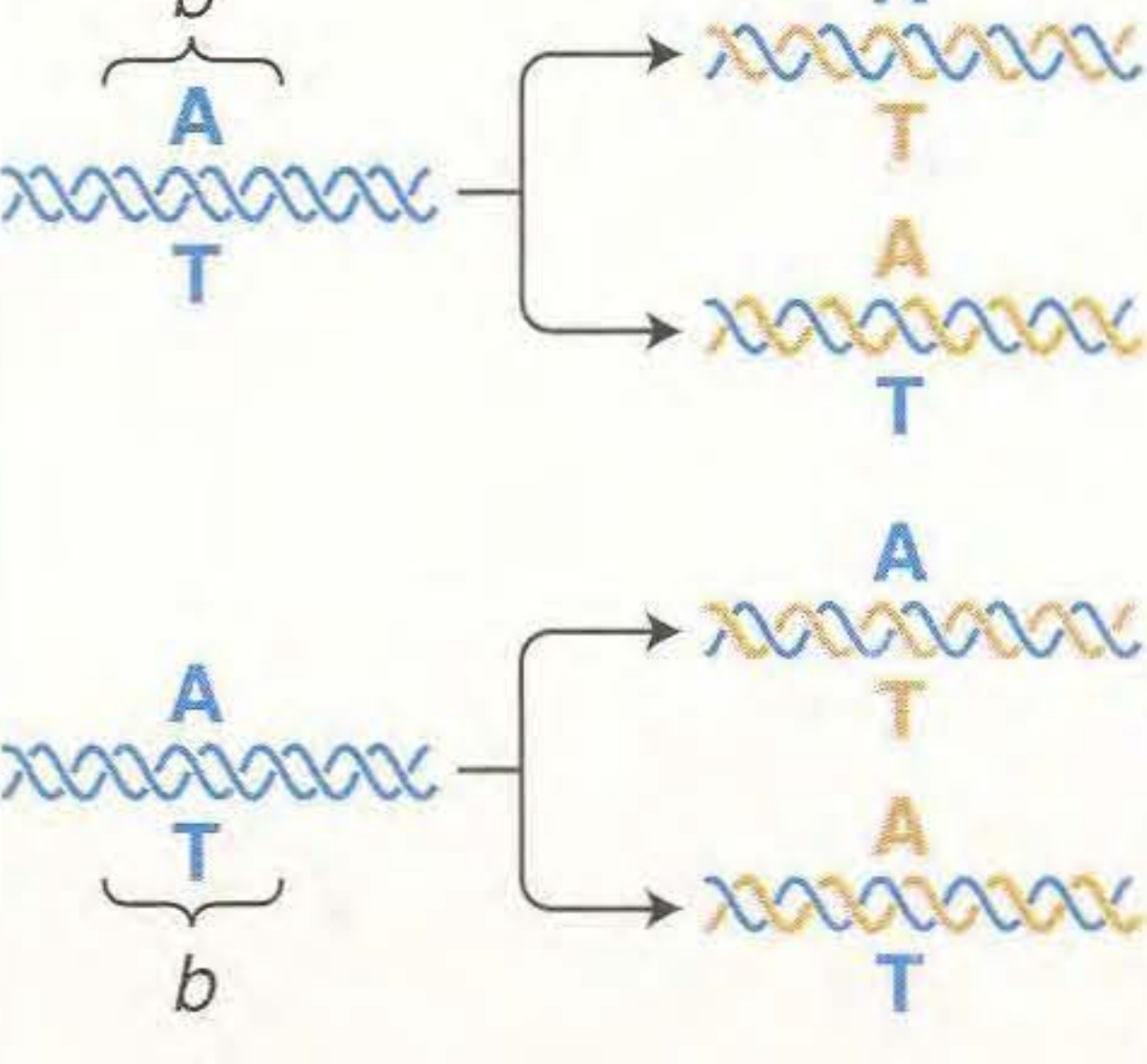
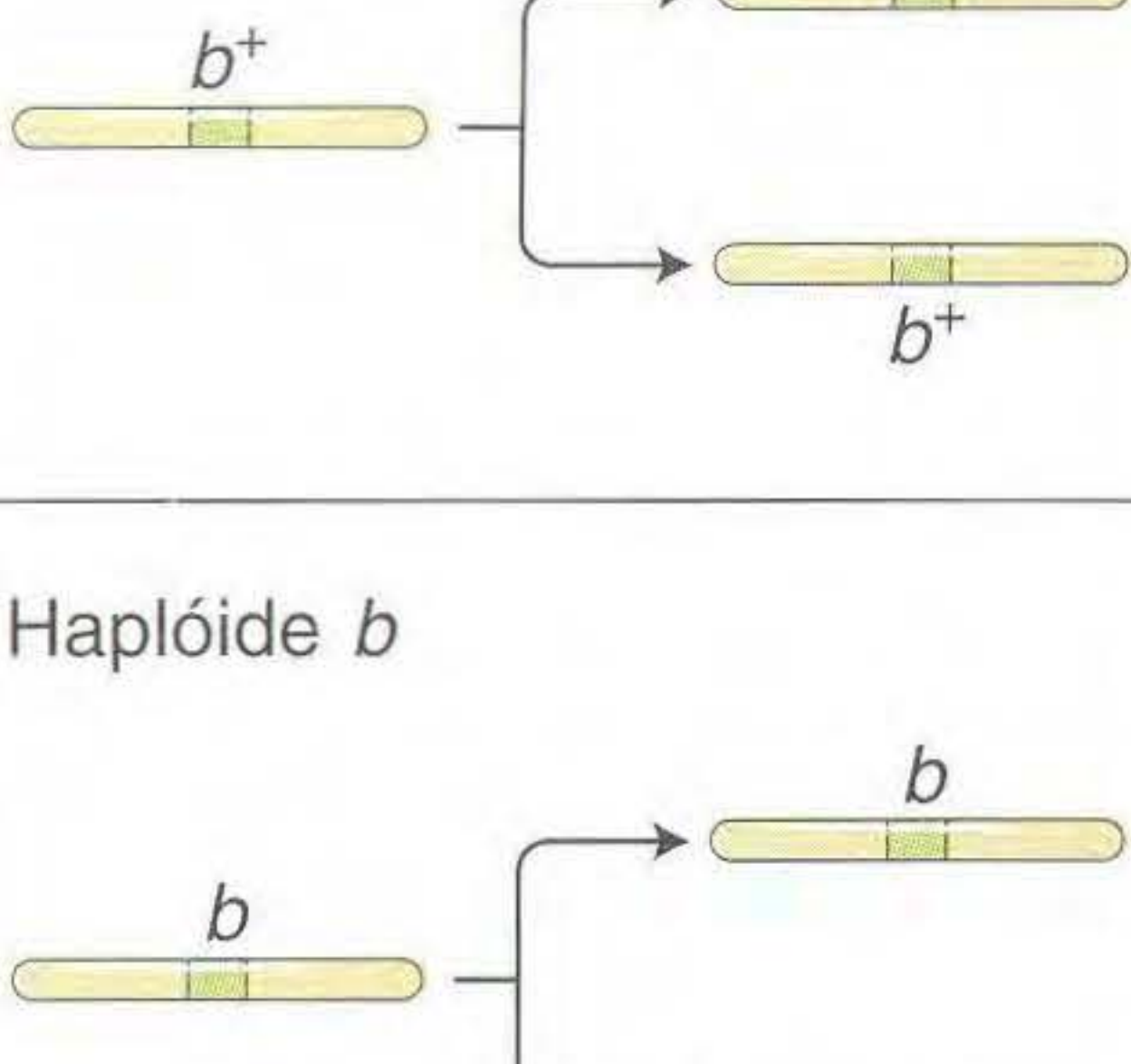
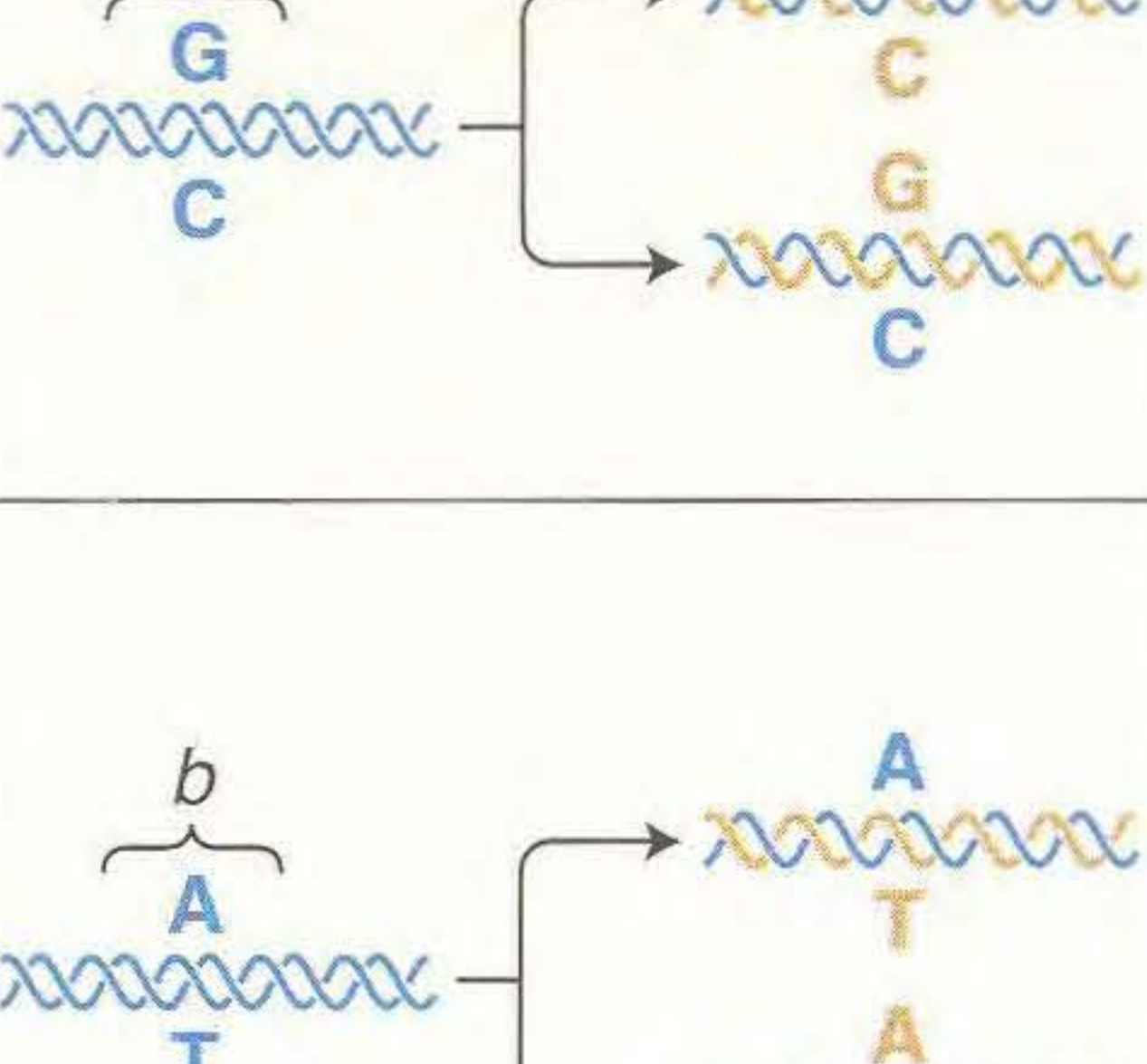
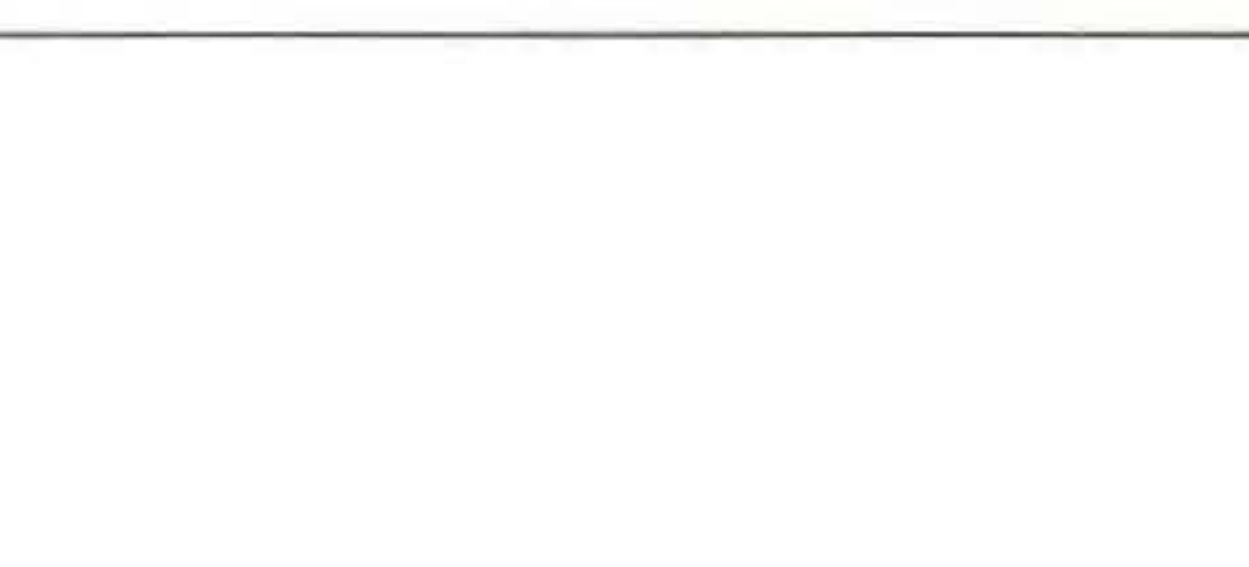

10 Intérfase



9 Telófase I

A célula divide-se

Moléculas de DNA replicam-se para formar cromátides idênticas

Formação de cromátide	Replicação do DNA
<p>Homozigoto diplóide b^+/b^+</p> 	
<p>Heterozigoto diplóide b^+/b</p> 	
<p>Homozigoto diplóide b/b</p> 	
<p>Haplóide b^+</p> 	
<p>Haplóide b</p> 	

A enzima PAH é feita de um único tipo de proteína. Que mudanças ocorreram na forma mutante do DNA do gene de PKU, e como tal mudança no DNA afeta o funcionamento da proteína e produz o fenótipo da doença? O sequenciamento dos alelos mutantes de muitos pacientes com PKU revelou muitas mutações em sítios diferentes ao longo do gene, principalmente nas regiões codificantes da proteína, os éxons. Os resultados estão resumidos na Figura 2.21. Eles representam uma gama de mudanças no DNA, mas a maioria são pequenas mudanças afetando apenas um par de nucleotídeos entre os milhares que constituem o gene. O que esses alelos têm em comum é que eles codificam uma proteína defeituosa que não tem mais a atividade normal de PAH. Mudando um ou mais aminoácidos, as mutações todas inativam alguma parte essencial da proteína codificada pelo gene. O efeito da mutação no funcionamento do gene depende de onde dentro do gene ocorre a mutação. Uma região funcional importante do gene é a que codifica o sítio de atividade da enzima; assim, essa região é muito sensível à mutação. Além disso, uma minoria de mutações são encontradas dentro de íntrons, e essas mutações em geral impedem o processamento normal do RNA transcrito primário. Algumas das consequências gerais da mutação ao nível de proteína são mostradas na Figura 2.22. Muitos dos alelos mutantes são de um tipo geralmente chamado de **alelos nulos**: as proteínas codificadas por eles não têm nenhuma função de PAH. Outros alelos mutantes reduzem o nível de funcionamento da enzima: às vezes eles são chamados de **mutações vazantes**, porque alguma função tipo selvagem parece “vazar” no fenótipo mutante.

A maioria das mutações altera a sequência de aminoácidos do produto proteico do gene, resultando em função reduzida ou ausente.

Temos perseguido a idéia de que encontrar um conjunto de genes que atuam na propriedade biológica investigada é uma meta importante da Genética, porque ela define os componentes do sistema. Entretanto, encontrar o modo *preciso* pelo qual os alelos mutantes levam a fenótipos mutantes geralmente é um desafio, que requer não só a identificação dos produtos proteicos desses genes, mas também estudos celulares e fisiológicos detalhados para medir os efeitos das mutações. Além disso, descobrir como conjuntos de genes *interagem* é um segundo nível de desafio e um tópico que será perseguido depois, começando no Capítulo 6.

FIG. 2.18 Cada cromossomo divide-se longitudinalmente em duas cromátides (*esquerda*); a nível molecular (*direita*), a única molécula de DNA de cada cromossomo replica-se, produzindo duas moléculas de DNA, uma para cada cromátide. Também são mostradas várias combinações de um gene com o alelo tipo selvagem b^+ e uma forma mutante b , causada pela mudança em um único par de bases de GC para AT. Note que, ao nível do DNA, as duas cromátides produzidas quando um cromossomo se replica são sempre idênticas uma à outra e ao cromossomo original.

Divisão nuclear ao nível de DNA

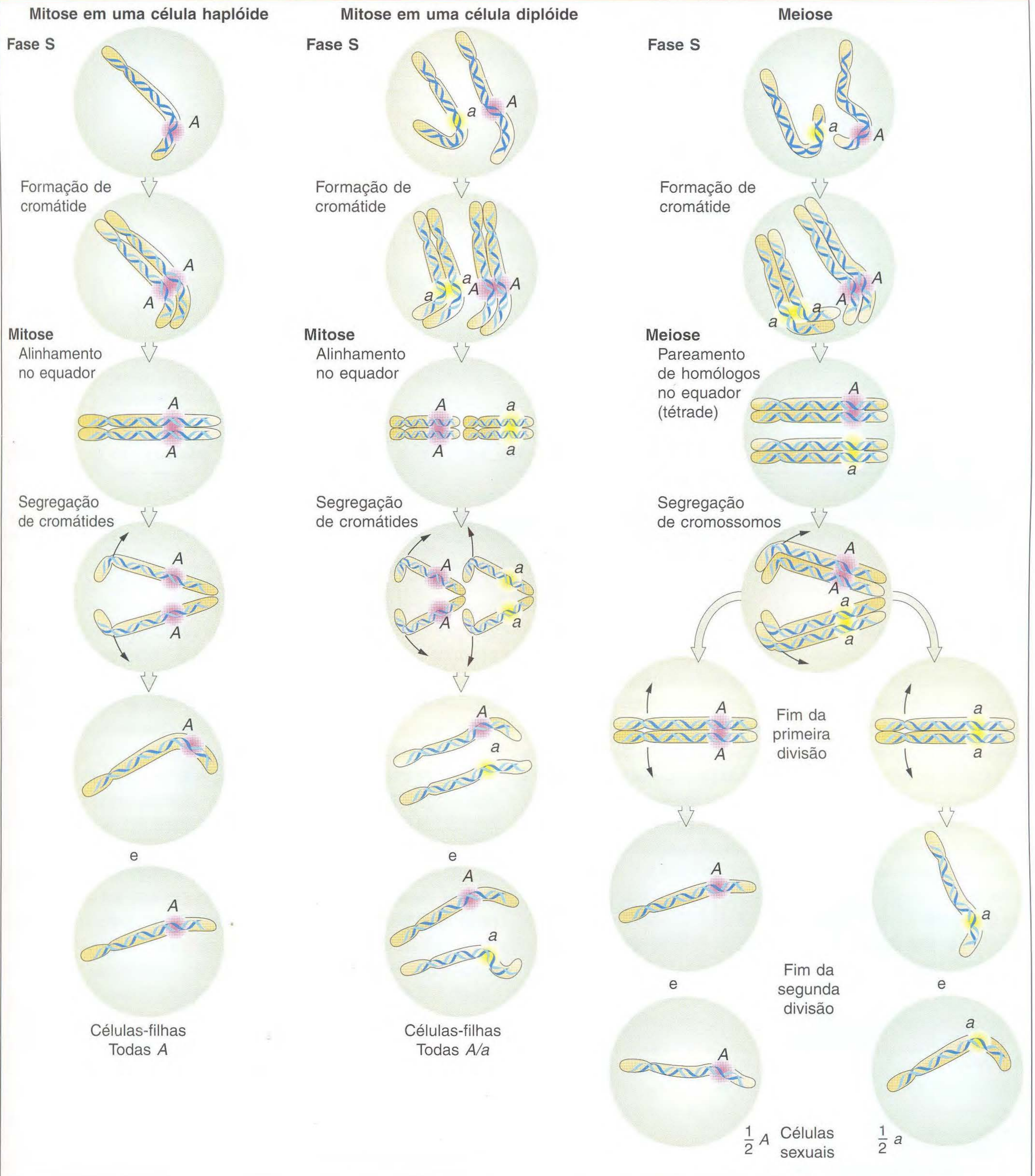


FIG. 2.19 DNA e transmissão gênica na mitose e meiose em eucariontes. São mostrados a fase S e os principais estágios da mitose e meiose. A divisão mitótica (*esquerda e meio*) conserva o genótipo da célula original. À direita, as duas divisões meióticas sucessivas que ocorrem durante o estágio sexual do ciclo vital têm o efeito de reduzir à metade o número de cromossomos. Os alelos *A* e *a* de um gene são usados para mostrar como os genótipos são transmitidos na divisão celular.

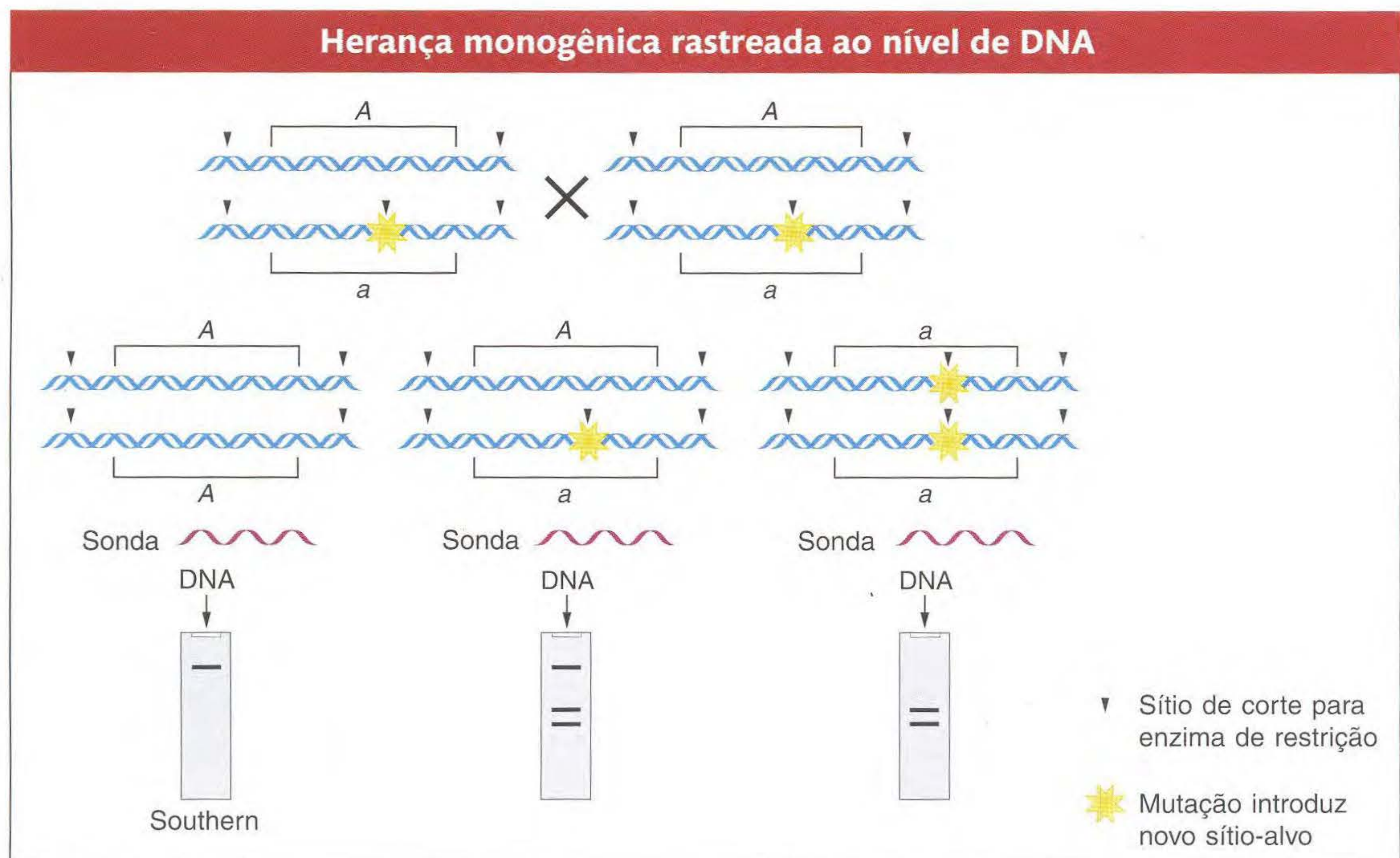


FIG. 2.20 Uma mutação recessiva que produz um alelo *a* por acaso também introduz um novo sítio de corte para uma enzima de restrição. Esse sítio de corte permite que a herança da mutação seja rastreada usando uma análise de Southern. A transferência de Southern detecta um fragmento de DNA em indivíduos homozigotos normais (*A/A*) e dois fragmentos em indivíduos albinos (*a/a*), mas ela detecta três fragmentos em indivíduos heterozigotos, devido à presença de alelos mutantes e normais.

Dominância e recessividade Com uma compreensão de como os genes funcionam através de seus produtos proteicos, podemos compreender melhor a dominância e a recessividade. A dominância já foi definida antes neste capítulo como o fenótipo mostrado pelo heterozigoto. Assim, formalmente, é o *fenótipo* que é dominante ou recessivo, mas, na prática, os geneticistas aplicam o termo mais frequentemente aos alelos. Essa definição formal não tem conteúdo molecular, mas tanto a dominância quanto a recessividade podem ter explicações simples a nível molecular. Introduzimos o tópico aqui, mas ele será revisto no Capítulo 6. A recessividade é observada nas mutações nos genes que são funcionalmente **haplossuficientes**. Embora uma célula diplóide normalmente tenha duas cópias tipo selvagem de

um gene, uma cópia de um gene haplossuficiente fornece produto gênico suficiente (em geral, uma proteína) para efetuar as reações normais da célula. Em um heterozigoto (digamos, $+/m$ onde m é nulo), a cópia restante codificada pelo alelo $+$ fornece produto proteico suficiente para a função normal.

Outros genes são **haploinsuficientes**. Em tais casos, um alelo mutante nulo será dominante porque, em um heterozigoto ($+/M$), o único alelo tipo selvagem não pode dar produto suficiente para a função normal.

Em alguns casos, as mutações resultam em uma *nova função* para o gene. Tais mutações podem ser dominantes porque, em um heterozigoto, o alelo tipo selvagem não pode mascarar essa nova função.



FIG. 2.21 Muitas mutações do gene de fenilalanina hidroxilase humana que causam o defeito da enzima são conhecidas. O número de mutações nos éxons, ou regiões codificantes de proteína (preto), é citado acima do gene. O número de mutações nas regiões dos íntrons (verde, número 1 até 13) que alteram a recomposição é citado abaixo do gene. [De C.R. Scriver, *Ann. Rev. Genet.* 28, 1994, 141-165.]

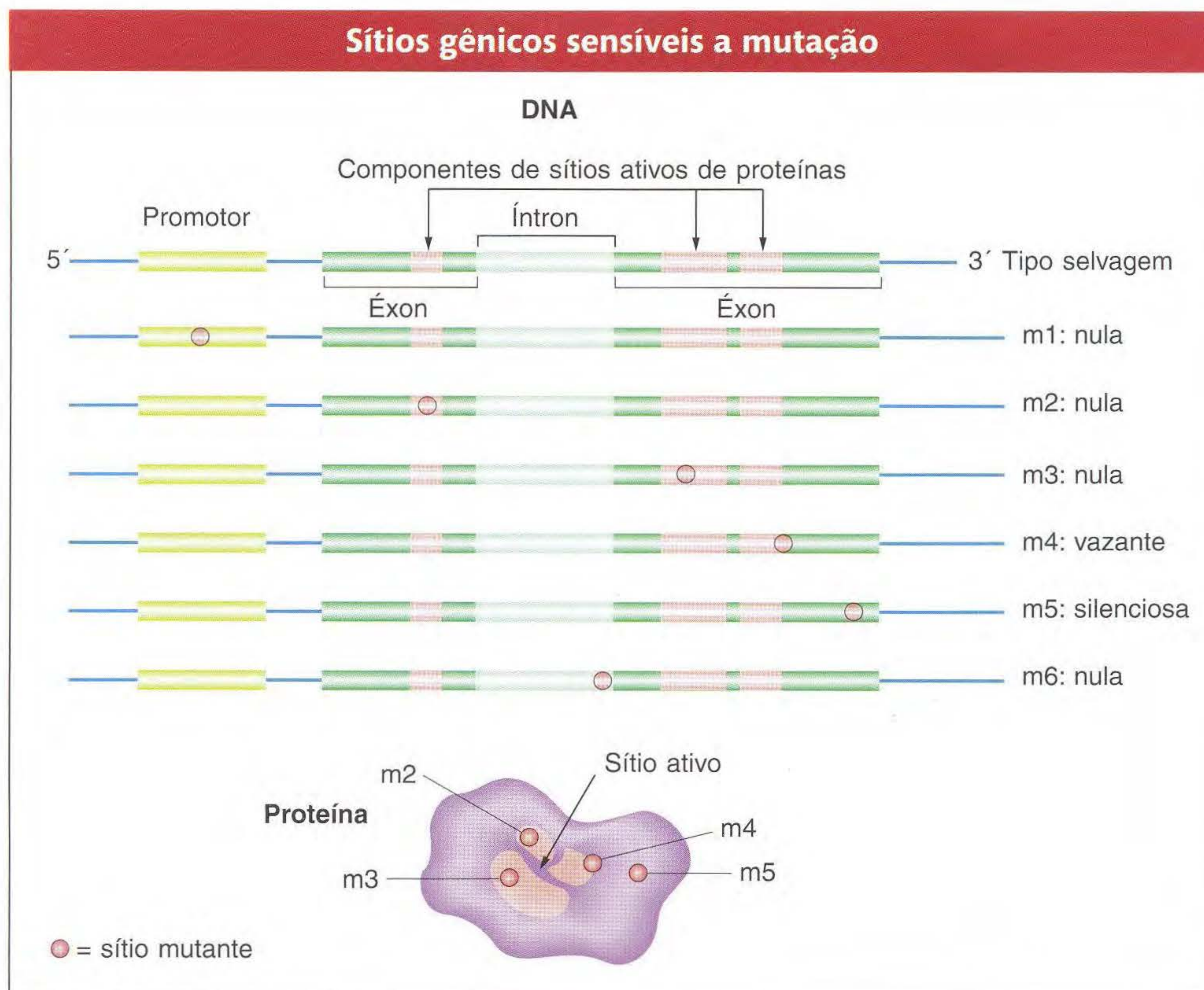


FIG. 2.22 Mutações nas partes de um gene codificando sítios ativos de enzimas levam a enzimas que não funcionam (mutações nulas). As mutações em outras partes do gene podem não ter efeito no funcionamento da enzima (mutações silenciosas).

A força molecular da segregação Mendel identificou a segregação como um processo hipotético, sem nenhuma compreensão de como ela funciona. Entretanto, agora compreendemos que o fuso nuclear é a máquina molecular que fornece a força motriz que separa os cromossomos ou cromátides tanto na mitose quanto na meiose. A Figura 2.23 ilustra a mitose como um exemplo. Na meiose em um heterozigoto, o fuso é o mecanismo que resulta na segregação alélica. Em qualquer divisão nuclear, as fibras do fuso formam-se de modo paralelo ao eixo da célula. As fibras do fuso são polímeros de uma proteína chamada tubulina. Elas se prolongam dos pólos da célula para os cromossomos dispostos no equador da célula. Cada centrômero atua como um sítio ao qual se liga um complexo multiproteico chamado **cinetócoro**. O cinetócoro atua como um sítio para a ligação aos microtúbulos das fibras do fuso. Os microtúbulos (variando em número de um a muitos) de um pólo ligam-se a um cinetócoro, e um número similar de microtúbulos do pólo oposto ligam-se ao cinetócoro da cromátide homóloga (mitose) ou cromossomo (meiose). Essas fibras arrastam os cromossomos para os pólos, onde eles formam núcleos-filhos. Embora as fibras do fuso pareçam cordas, sua ação não é similar à de uma corda. Em vez disso, a tubulina se despolimeriza nos cinetócoros, encurtando o microtúbulo e, assim, exercendo uma força de puxar. Depois, as unidades segregantes são mais separadas por proteínas moleculares que atuam em outro conjunto de microtúbulos não-conectados ao cinetócoro, mas também correndo de pólo a pólo. O aparelho do fuso e o

complexo de cinetócoros e centrômeros juntos determinam a fidelidade da divisão nuclear. É marcante que, na maioria das divisões celulares, as fibras do fuso separam com precisão o DNA. Entretanto, todas as máquinas podem falhar, e a falha ocasional das fibras do fuso leva a núcleos com mais ou menos o número normal de cromossomos, e essa falha leva a condições específicas tais como o distúrbio humano síndrome de Down, como veremos no Capítulo 16.

► 2.4 Descobrindo Genes pela Observação das Proporções de Segregação

Lembre que um dos objetivos gerais da análise genética hoje em dia é dissecar uma propriedade biológica descobrindo o conjunto de genes isolados que a afetam. Aprendemos que um modo importante de identificar esses genes é pelas proporções fenotípicas de segregação geradas por suas mutações, a maioria delas 1:1 e 3:1, ambas as quais são baseadas na segregação igual como definido por Gregor Mendel.

Vejamos alguns exemplos que ampliam o enfoque mendeliano para uma situação experimental moderna. Tipicamente, o pesquisador é confrontado com uma gama de fenótipos mutantes interessantes que afetam a propriedade de interesse (tal como as mostradas na Figura 2.1) e agora precisa saber se eles são herdados como alelos mutantes isolados. Alelos mutantes podem ser dominantes ou recessivos, depen-

Ação dos microtúbulos na mitose

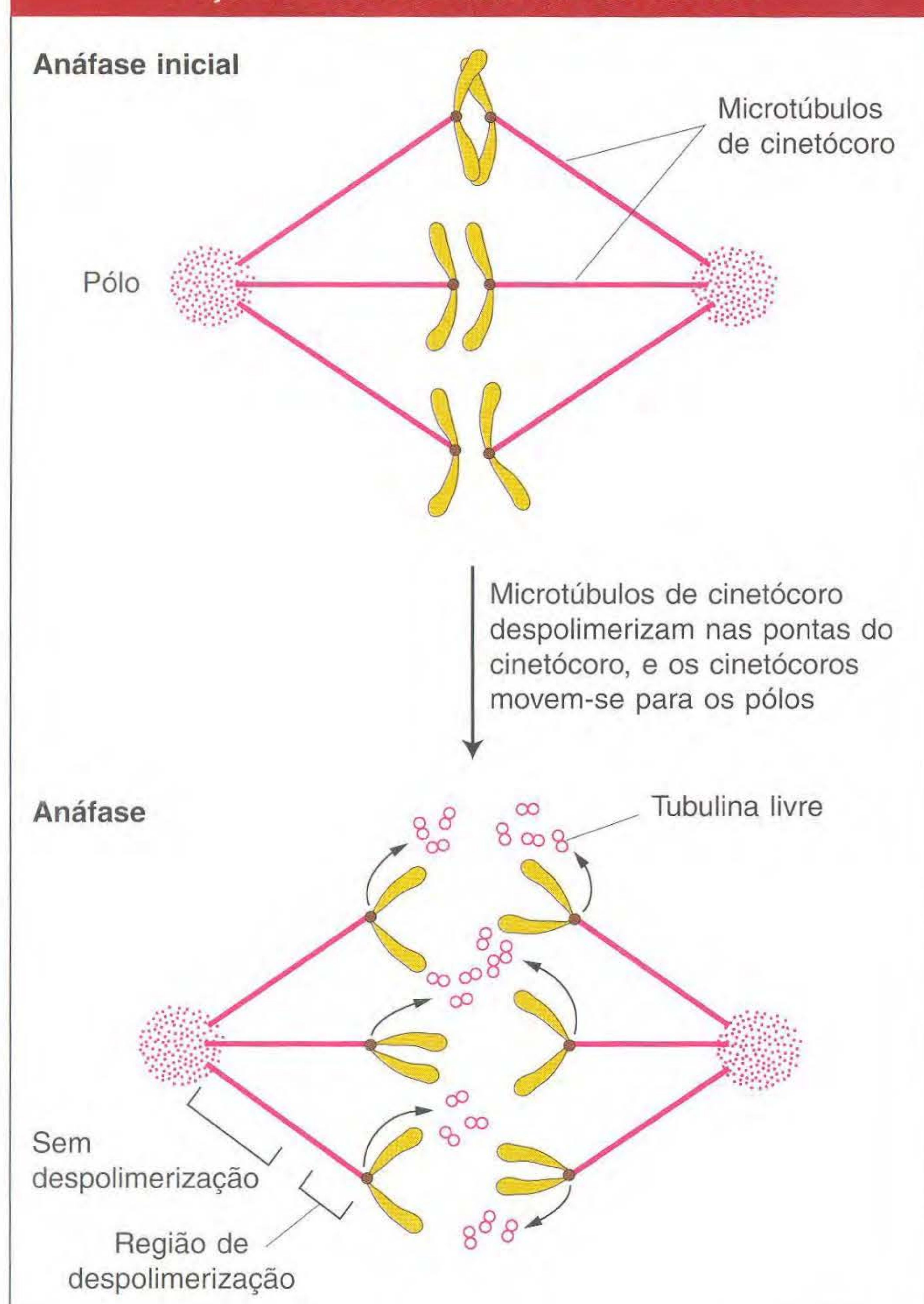


FIG. 2.23 Os microtúbulos tracionam as cromátides despolarizando em subunidades de tubulinas nos cinetócoros. [De G.J. Gorbsky, P.J. Sammak, and G. Borisy, *J. Cell Biol.* 104, 1987, 9; e G.J. Gorbsky, P.J. Sammak, and G. Borisy, *J. Cell Biol.* 106, 1988, 1185; modificado de H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell, *Molecular Cell Biology*, 4th ed. Copyright 2000 by W.H. Freeman and Company.]

dendo de sua ação. Assim, a questão da dominância também precisa ser considerada na análise.

O procedimento padrão é cruzar o mutante com o tipo selvagem. (Se o mutante for estéril, então será necessário um outro enfoque.) Primeiro consideremos três casos simples que cobrem a maioria dos resultados possíveis:

1. Uma flor mutante fértil sem pigmento nas pétalas (por exemplo, pétalas brancas em contraste com as vermelhas normais)
2. Uma mosca-das-frutas fértil mutante com asas pequenas
3. Um fungo mutante fértil que não produz conídios (esporos assexuais)

Descobrimo um gene ativo no desenvolvimento da cor da flor

Para começar o processo, a planta com flores brancas é cruzada com o tipo selvagem normal vermelho. Todas as plantas de F_1 têm flores vermelhas, e, das 500 plantas de F_2 da amos-

tra, 378 têm flores vermelhas e 122 têm flores brancas. Se reconhecermos a existência de um erro de amostragem, esses números de F_2 estão muito próximos de uma proporção de $\frac{3}{4} : \frac{1}{4}$, ou de 3:1. Como essa proporção indica herança monogênica, podemos concluir que o mutante é causado por uma alteração recessiva em um único gene. De acordo com as regras gerais de nomenclatura gênica, o alelo mutante para pétalas brancas pode ser chamado de *alb* para albino, e o alelo tipo selvagem seria *alb*⁺ ou simplesmente +. (As convenções para nomenclatura dos alelos variam um pouco entre os organismos: algumas das variações são mostradas no apêndice sobre nomenclatura.) Concluimos que o alelo tipo selvagem tem um papel essencial na produção de pétalas coloridas da planta, uma propriedade que é quase certamente necessária para atrair polinizadores para a flor. O gene pode estar implicado na síntese bioquímica do pigmento ou na parte de sistema de sinalização que diz às células da flor para começar a fazer pigmento, ou em várias outras possibilidades que necessitam de mais investigação. A nível puramente genético, os cruzamentos feitos seriam simbolicamente representados como

P	+/+ × alb/alb
F ₁	todos +/alb
F ₂	$\frac{1}{4}$ +/+
	$\frac{1}{2}$ +/alb
	$\frac{1}{4}$ alb/alb

Descobrimo um gene para desenvolvimento das asas

No exemplo da mosca-das-frutas, o cruzamento entre uma mosca mutante de asas curtas com uma selvagem de asas longas do estoque deu uma prole de 788, classificada do seguinte modo:

196 machos de asas curtas
194 fêmeas de asas curtas
197 machos de asas longas
201 fêmeas de asas longas

No total, existem 390 de asas curtas e 398 de asas longas na prole, muito próximo da proporção 1:1. A proporção é a mesma dentro dos machos e fêmeas, novamente dentro dos limites de erro de amostragem. Assim, a partir desses resultados, o mutante “asas curtas” era muito provavelmente produzido por uma mutação dominante. Note que, para uma mutação dominante ser expressa, é necessária apenas uma “dose” do alelo mutante; assim, na maioria dos casos, quando o mutante primeiro se apresenta na população, ele estará no estado heterozigoto. (Isso não é verdade para uma mutação recessiva tal como a do exemplo anterior de plantas, que tem que ser homozigota para ser expressa e deve ter vindo da autofecundação de uma planta heterozigota não-identificada na geração anterior.)

Quando a prole de asas longas foi entrecruzada, todas suas descendentes eram de asas longas, como esperado de um alelo tipo selvagem recessivo. Quando a prole de asas curtas foi entrecruzada, seus descendentes mostraram uma

proporção de três quartos de asas curtas para um quarto de asas longas.

As mutações dominantes são representadas por letras maiúsculas ou palavras: no presente exemplo, o alelo mutante pode ser chamado *SH*, de curta. Então, os cruzamentos seriam simbolizados por:

$$\begin{array}{lcl} \text{P} & +/+ & \times \text{SH}/+ \\ \text{F}_1 & \frac{1}{2} +/+ & \\ & \frac{1}{2} \text{SH}/+ & \\ & +/+ & \times +/+ \\ \text{F}_2 & \text{todos } +/+ & \\ & \text{SH}/+ & \times \text{SH}/+ \\ \text{F}_2 & \frac{1}{4} \text{SH}/\text{SH} & \\ & \frac{1}{2} \text{SH}/+ & \\ & \frac{1}{4} +/+ & \end{array}$$

Essa análise do mutante da mosca identifica um gene que é parte de um subgrupo de genes que, na forma tipo selvagem, são cruciais para o desenvolvimento normal de uma asa. Tal resultado é o ponto de partida para mais estudos que enfocariam o exato desenvolvimento e modos celulares nos quais o crescimento da asa é parado, que, identificados, revelariam o tempo de ação do alelo tipo selvagem no curso do desenvolvimento.

Descobrimo um gene para a produção de esporos

O fungo sem esporos foi cruzado com um fungo tipo selvagem tendo esporos. Em uma amostra de 300 da prole, 152 tinham esporos e 148 não os tinham, muito perto de uma proporção 1:1. Deduzimos dessa proporção monogênica que a mutação sem esporos é de um único gene. Nos haplóides, atribuir a dominância geralmente não é possível, mas, por conveniência, podemos chamar o alelo sem esporos como *sp* e o alelo tipo selvagem como *sp*⁺ ou +. O cruzamento deve ter sido

$$\begin{array}{lcl} \text{P} & + & \times \text{sp} \\ \text{Meiótico diplóide} & +/\text{sp} & \\ \text{F}_1 & \frac{1}{2} + & \\ & \frac{1}{2} \text{sp} & \end{array}$$

A mutação e a análise da herança descobriram um gene cujo tipo selvagem é essencial para sintetizar esporos, o principal agente de dispersão de um fungo. Os esporos são normalmente produzidos por um processo de podar nas pontas de hifas aéreas especiais. Agora o mutante precisa ser investigado para ver a posição na seqüência desenvolvimental na qual o mutante produz um bloqueio. Essa informação revelará a época e o local nas células nas quais atua o alelo normal.

Às vezes, a gravidade de um fenótipo mutante torna o organismo estéril, incapaz de prosseguir o ciclo sexual. Como pode ser demonstrada a herança monogênica de mutantes

estéreis? Em um organismo diplóide, um mutante recessivo estéril pode ser propagado como um heterozigoto, e, então, o heterozigoto pode ser autofecundado para produzir os 25% de mutantes esperados para estudo. Um mutante dominante estéril é um beco sem saída genético e não pode ser propagado sexualmente, mas, nas plantas e fungos, tal mutante pode ser facilmente propagado assexualmente.

E se o cruzamento entre um mutante e um tipo selvagem não produzir uma proporção 3:1 ou 1:1 como discutido aqui, mas uma outra proporção? Tal resultado pode ser devido a interações de vários genes ou a um efeito ambiental. Algumas dessas possibilidades são discutidas no Capítulo 6.

Os resultados da descoberta do gene

Voltemos às fotos na Figura 2.1 para ver como a descoberta do gene pode levar a um conjunto de genes cuja interação pode dar modelos de importantes mecanismos celulares.

Um conjunto de genes instrumentais no desenvolvimento de flores Todas as formas anormais das flores *Arabidopsis* mostradas na Figura 2.1 são herdadas como mutações monogênicas. Esses mutantes e outros como eles têm sido fundamentais para compreender a importante propriedade biológica do desenvolvimento floral. Todos eles são exemplos de mutantes *homeóticos*, definidos como aqueles nos quais o alelo mutante tem o efeito de transformar um tipo de órgão em outro. Os mutantes homeóticos foram encontrados também em animais. No caso de *Arabidopsis*, a flor tipo selvagem tem quatro verticilos de órgãos, que podem ser representados (começando da parte de fora da flor) do seguinte modo:

tipo selvagem (WT): sépalas, pétalas, estames, carpelos

Todos os mutantes têm arranjos de verticilos anormais do seguinte modo:

apetala 1 e 2 (ap1 e ap2): carpelos, estames, estames, carpelos

apetala 3 (ap3): sépalas, sépalas, carpelos, carpelos

agamous (ag): repetições da unidade “sépalas, pétalas, pétalas”

(Assim, *agamous* é estéril porque não tem órgãos reprodutivos, os estames ou carpelos.)

A descoberta, a partir de tais mutantes, de que órgãos são interconversíveis é muito interessante porque mostra que as plantas estão em um limite desenvolvimental que pode ir em várias direções. Os estudos de mutantes precedentes revelaram que todos os genes identificados codificam vários fatores de transcrição, que formam gradientes superpostos no meristema floral, onde seu efeito combinado promove um desenvolvimento de um órgão específico.

Vejamos um exemplo de como funciona o modelo. As pétalas são determinadas no meristema em desenvolvimento da flor pela superposição dos gradientes de dois fatores geralmente chamados A e B. O gradiente A em si (em regiões não-superpostas) determina sépalas. O gene *apetala 1* atua

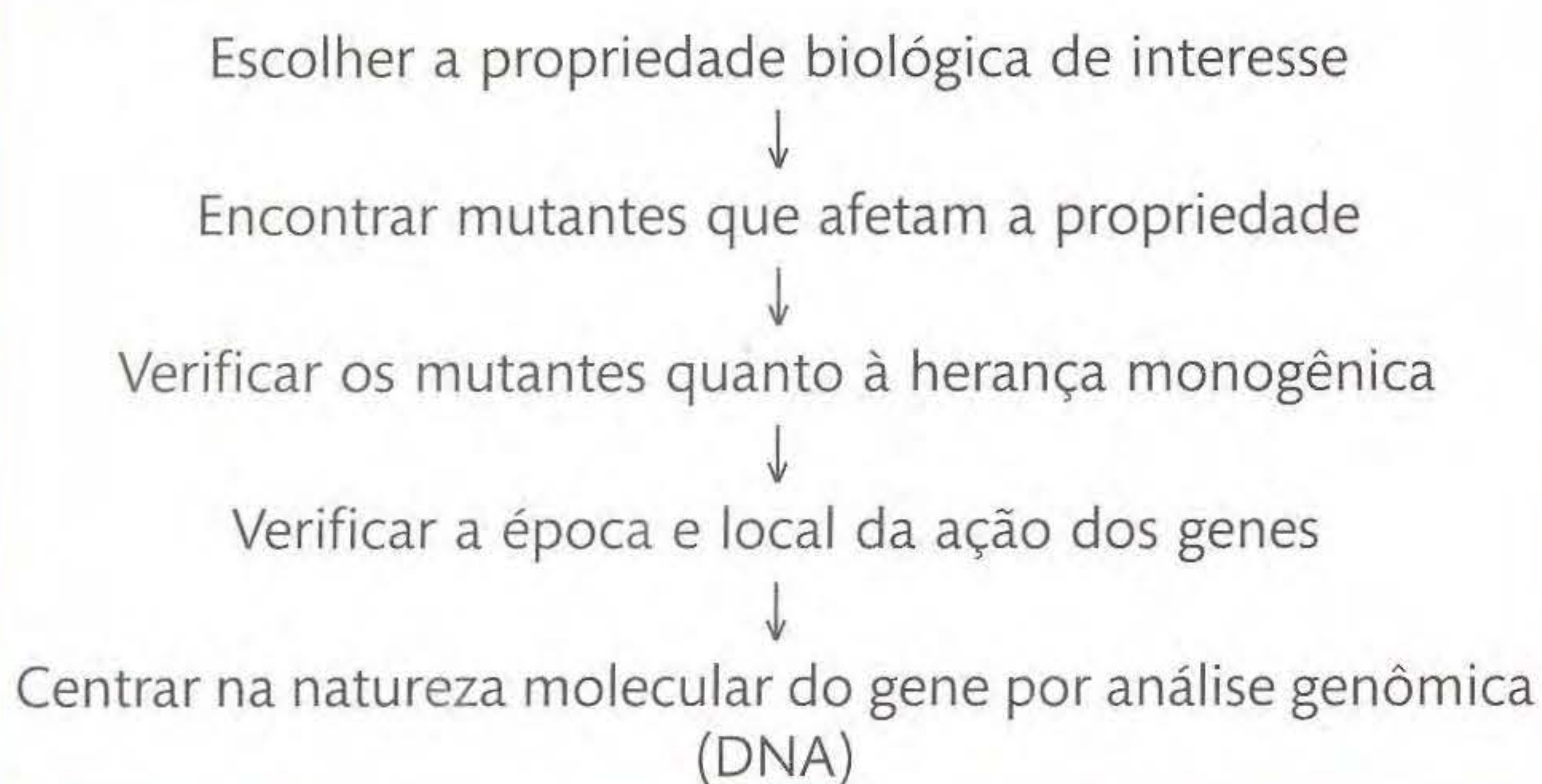
para construir o gradiente A; logo, em um mutante *ap1* nem sépalas nem pétalas são produzidas, e a flor contém apenas estames e carpelos. Gradientes similares explicam todos os outros fenótipos mutantes.

Um conjunto de genes instrumentais no desenvolvimento de hifas Entre os mutantes de fungo do tipo mostrado na Figura 2.1 estão aqueles com defeitos nas proteínas do citoesqueleto actina e tubulina; nas proteínas que fazem parte da estrutura e mantêm os gradientes intracelulares de cálcio necessários para o crescimento normal; nas enzimas que afrouxam as paredes celulares na ponta; e nas enzimas chamadas cinases que ativam ou inativam outras proteínas durante a transdução intracelular de sinal. No geral, esses achados revelam o subgrupo de genes que codificam os componentes celulares que interagem para produzir o crescimento normal de hifas.

Genética direta

Em geral, o tipo de enfoque para a descoberta do gene que temos seguido é, às vezes, chamado de **genética direta**, um enfoque para a compreensão da função biológica que começa com mutantes monogênicos aleatórios e termina com a análise bioquímica e celular detalhada deles, em geral incluindo a análise genômica. (Veremos a **genética reversa** em funcionamento nos capítulos posteriores. Em resumo, ela começa com a análise genômica para identificar um conjunto de genes como candidatos para codificar a propriedade biológica de interesse, e, então, inclui mutantes direcionados especificamente para esses genes, examinando os fenótipos mutantes para ver se eles afetam a propriedade em estudo.)

A descoberta do gene pela herança monogênica é, às vezes, chamada de genética direta. Em geral, ela funciona na seguinte seqüência:



Prever as proporções da prole ou genótipos parentais aplicando os princípios da herança monogênica

Podemos resumir a direção da análise da descoberta do gene do seguinte modo:

Observar as proporções fenotípicas na prole →
Deduzir os genótipos dos genitores (A/A , A/a ou a/a)

Entretanto, o mesmo princípio de herança (essencialmente a lei de Mendel da segregação igual) também pode

ser usado para prever as proporções fenotípicas na prole de genitores de *genótipos conhecidos*. Esses genitores seriam de estoques mantidos pelo pesquisador. Os tipos e proporções da prole de cruzamentos tais como $A/A \times A/a$, $A/A \times aa$, $A/a \times A/a$ e $Aa \times a/a$ podem ser facilmente previstos. Em resumo:

Cruzar genitores de genótipos conhecidos →

Prever proporções fenotípicas na prole

Esse tipo de análise é usado em cruzamentos gerais para sintetizar genótipos para pesquisa ou para agricultura. Ele também é útil para prever probabilidades de vários resultados em reproduções humanas em famílias com histórias de doenças monogênicas.

Após a herança monogênica ter sido estabelecida, um indivíduo apresentando o fenótipo dominante mas de *genótipo desconhecido* pode ser testado para ver se o genótipo é homozigoto ou heterozigoto. Tal teste pode ser feito cruzando o indivíduo (com o fenótipo $A/?$) com uma linhagem testadora recessiva a/a . Se o indivíduo é heterozigoto, resultará uma proporção 1:1 ($\frac{1}{2} A/a$ e $\frac{1}{2} a/a$); se o indivíduo é homozigoto, toda a prole apresentará o fenótipo dominante (todos A/a). Em geral, o cruzamento de um indivíduo de heterozigose desconhecida (para um gene ou mais) com um genitor totalmente recessivo é chamado de **cruzamento-teste**, e o indivíduo recessivo é chamado **testador**. Encontraremos cruzamentos-teste muitas vezes nos capítulos subseqüentes; eles são muito úteis em deduzir os eventos meióticos tomando parte em genótipos mais complexos, tais como diíbridos e triíbridos. O uso de um testador totalmente recessivo significa que a meiose no genitor testador pode ser ignorada porque todos os seus gametas são recessivos e não contribuem para os fenótipos da prole. Um teste alternativo para heterozigose (útil se o testador recessivo não estiver disponível e o organismo puder ser autofecundado) é simplesmente autofecundar o desconhecido: se o organismo sendo testado é heterozigoto, será encontrada uma proporção 3:1 na prole. Tais testes são úteis e comuns na análise genética rotineira.

Os princípios da herança (tais como a lei da segregação igual) podem ser aplicados em duas direções: (1) deduzir genótipos a partir de proporções fenotípicas e (2) prever proporções fenotípicas a partir de genitores com genótipos conhecidos.

► 2.5 Padrões de Herança Monogênica Ligada ao Sexo

Os cromossomos que analisamos até agora são autossomos, os cromossomos “regulares” que formam a maioria do conjunto genômico. Entretanto, muitos animais e plantas têm um par especial de cromossomos associados ao sexo. Os cromossomos sexuais também segregam igualmente, mas as proporções fenotípicas vistas na prole são, em geral, diferentes das proporções autossômicas.

Cromossomos sexuais

A maioria dos animais e muitas plantas apresentam dimorfismo sexual; em outras palavras, os indivíduos são masculinos ou femininos. Na maioria desses casos, o sexo é determinado por um par especial de **cromossomos sexuais**. Vejamos os humanos como um exemplo. As células do corpo humano têm 46 cromossomos: 22 pares de homólogos autossomos mais 2 cromossomos sexuais. As mulheres têm um par de cromossomos sexuais idênticos chamados de **cromossomos X**. Os homens têm um par não-idêntico, consistindo em um X e um Y. O **cromossomo Y** é consideravelmente menor que o X. Assim, se representarmos os cromossomos autossômicos por A, podemos escrever

$$\text{mulheres} = 44A + XX$$

$$\text{homens} = 44A + XY$$

Na meiose nas mulheres, os dois cromossomos X do par segregam como os autossomos, e cada ovócito recebe um cromossomo X. Assim, com relação aos cromossomos sexuais, os gametas são apenas de um tipo e a mulher é dita como sendo o **sexo homogamético**. Na meiose nos homens, os cromossomos X e Y fazem par por uma região curta, que garante que o X e Y se separem de modo que há dois tipos de espermatozoides, metade com um X e a outra metade com um Y. Portanto, o homem é chamado de **sexo heterogamético**.

Os padrões de herança de genes nos cromossomos sexuais são diferentes dos genes autossômicos. Os padrões de herança de cromossomos sexuais foram primeiro investigados no início dos anos de 1900 no laboratório do grande geneticista Thomas Hunt Morgan, usando a mosca-das-frutas *Drosophila melanogaster* (veja boxe Organismo-modelo adiante). Esse inseto tem sido um dos mais importantes organismos de pesquisa em genética; ele é pequeno, e seu ciclo de vida simples contribui para sua utilidade nesse aspecto. As moscas-das-frutas têm três pares de autossomos mais um par de cromossomos sexuais, novamente chamados de X e Y. Como nos mamíferos, as fêmeas de *Drosophila* têm a constituição XX e os machos são XY. Entretanto, o mecanismo de determinação do sexo em *Drosophila* difere do de mamíferos. Na *Drosophila*, o número de cromossomos X em relação ao de autossomos determina o sexo: dois X resultam em uma fêmea e um X resulta em um macho. Nos mamíferos, a presença do cromossomo Y determina um macho e a ausência de um Y determina uma fêmea. Entretanto, é importante notar que, a despeito dessa base um pouco diferente para determinação do sexo, os padrões de herança monogênica dos genes nos cromossomos sexuais são acentuadamente similares em *Drosophila* e mamíferos.

As plantas vasculares apresentam uma variedade de arranjos sexuais. As espécies **dióicas** são as que apresentam dimorfismo sexual similar ao dos animais, com as plantas femininas tendo flores contendo apenas ovários e as plantas masculinas tendo flores contendo apenas anteras (Figura 2.24). Algumas, mas não todas, as plantas dióicas têm um par de cromossomos não-idênticos associados (e quase certamente determinantes) do sexo da planta. Das espécies com cromossomos sexuais não-idênticos, uma grande proporção

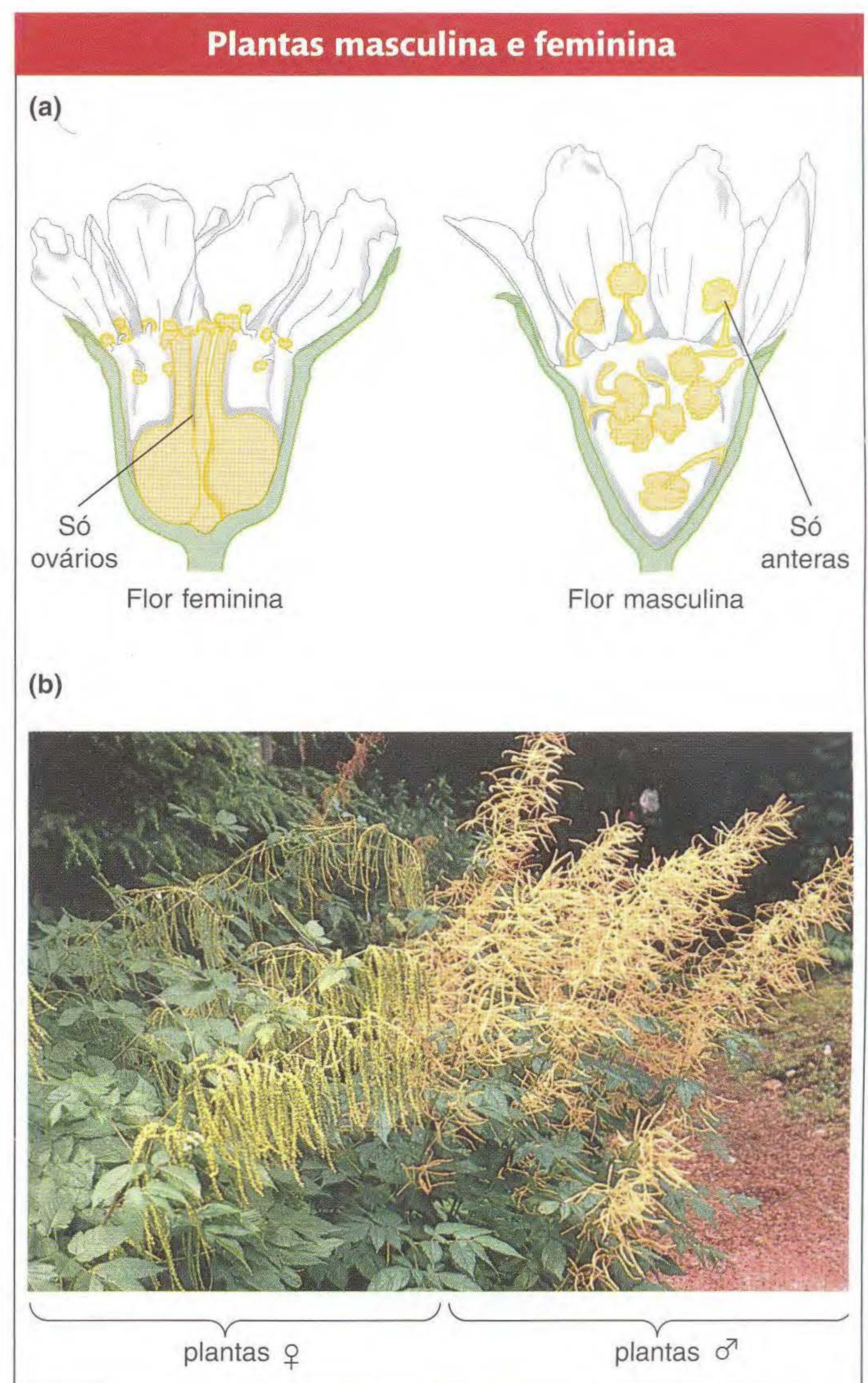


FIG. 2.24 Exemplos de duas plantas de espécies dióicas são (a) *Osmaronia dioica* e (b) *Aruncus dioicus*. [(a) Leslie Bohm; (b) Anthony Griffiths.]

tem um sistema XY. Por exemplo, a planta dióica *Melandrium album* tem 22 cromossomos por célula: 20 autossomos mais 2 cromossomos sexuais, com fêmeas XX e machos XY. Outras plantas dióicas não têm pares de cromossomos visivelmente diferentes de cromossomos; eles podem ainda ter cromossomos sexuais, mas estes não de tipos visivelmente distintos.

Padrões de herança ligada ao sexo

Os citogeneticistas dividem os cromossomos X e Y em regiões homólogas e diferenciais. Novamente, vamos usar os humanos como exemplo (Figura 2.25). As regiões diferenciais, que contêm a maioria dos genes, não têm contrapartes no outro cromossomo sexual. Assim, nos homens, os genes nas regiões diferenciais são ditos **hemizigotos** ("metade zigotos"). A região diferencial do cromossomo X contém muitas centenas de genes; a maioria desses genes não tomam parte da função sexual, e eles influenciam uma grande faixa de propriedades humanas. O cromossomo Y contém apenas algumas dúzias



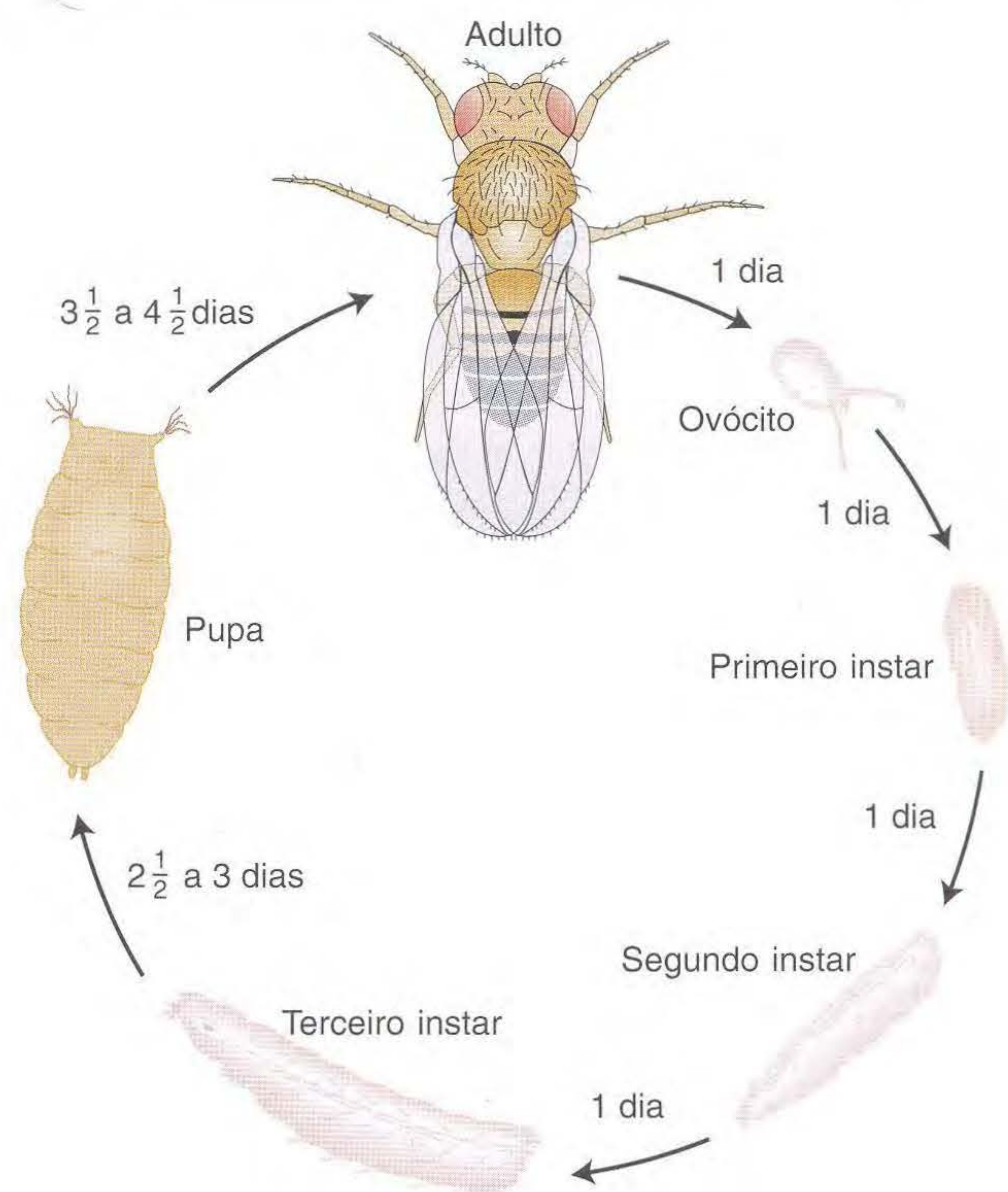
Organismo-modelo *Drosophila*

A *Drosophila melanogaster* foi um dos primeiros organismos-modelo a serem usados em genética. Ela está prontamente disponível em frutos maduros, tem um curto ciclo de vida e é simples de cultivar e cruzar. O sexo é determinado pelos cromossomos sexuais X e Y (XX = fêmea, XY = macho), e machos e fêmeas são facilmente distinguíveis. Os fenótipos mutantes surgem regularmente em populações de laboratório, e sua frequência pode ser aumentada por tratamento com radiação mutagênica ou substâncias químicas. É um organismo diplóide, com quatro pares de cromossomos homólogos ($2n = 8$). Nas glândulas salivares de alguns outros tecidos, várias rodadas de replicação de DNA sem divisão cromossômica resultam em “cromossomos gigantes”, cada um com um padrão único de bandejamento que fornece aos geneticistas marcos para o estudo de mapeamento cromossômico e rearranjos. Existem muitas espécies e raças de *Drosophila*, que têm sido matéria-prima importante para o estudo da evolução.

O tempo voa como uma flecha; as moscas-das-frutas gostam de bananas. (Grouxo Marx)



Drosophila melanogaster, a mosca-da-fruta. [SLP/Photo Researchers.]



Ciclo de vida de *Drosophila melanogaster*.

de genes. Alguns desses genes têm contrapartes no cromossomo X, mas a maioria não. Esse último tipo toma parte na função sexual masculina. Um desses genes, *SRY*, determina a própria masculinidade. Vários outros genes são específicos para a produção de espermatozóides nos machos.

Em geral, os genes nas regiões diferenciais são ditos apresentando padrões de herança chamados de **ligação ao sexo**. Os alelos mutantes na região diferencial do cromossomo X apresentam um padrão de herança monogênica chamado de **ligação ao X**. Os alelos mutantes dos poucos genes na região diferencial do cromossomo Y apresentam **ligação ao Y**. Um gene que é ligado ao sexo pode apresentar proporções fenotípicas que são diferentes em cada sexo. Quanto a isso, os padrões de herança ligada ao sexo contrastam com os padrões de herança de genes nos autossomos, que são os mesmos em cada sexo. Se a localização genômica de um gene é desconhecida, um padrão de herança ligada ao sexo indica que o gene está em um cromossomo sexual.

Os cromossomos humanos X e Y têm duas regiões homólogas curtas, uma em cada ponta (veja Figura 2.25). No sentido em que essas regiões são homólogas, elas são similares aos autossomos, e, portanto, são chamadas de **regiões pseudo-autossômicas 1 e 2**. Uma ou ambas essas regiões fazem par na meiose e sofrem crossing over (veja Capítulo 4 para deta-

lhes do crossing over). Por essa região, os cromossomos X e Y podem agir como um par e segregar em números iguais de espermatozóides.

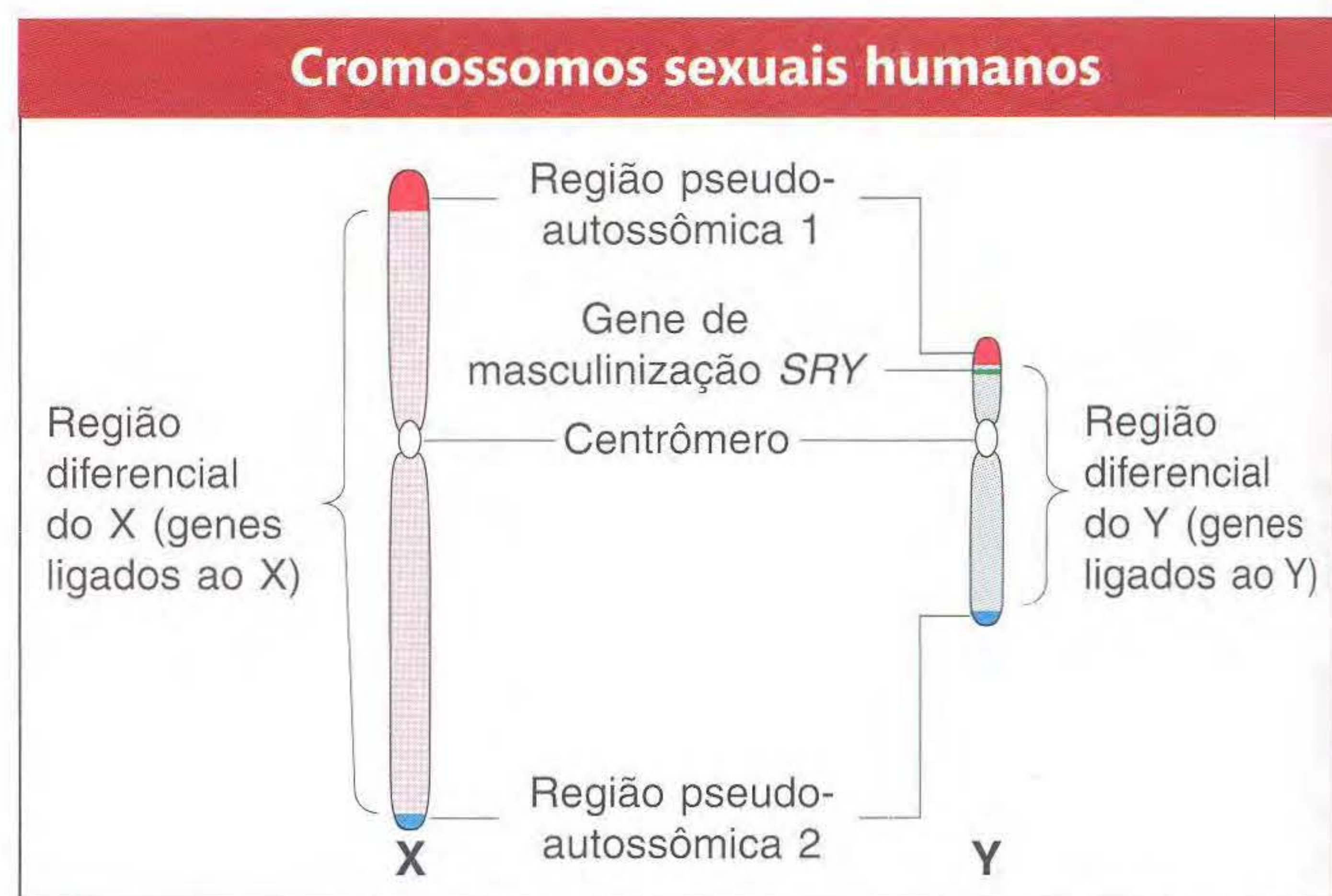


FIG. 2.25 Os cromossomos sexuais humanos contêm regiões diferenciais e duas regiões de pareamento. Essas regiões foram localizadas observando onde os cromossomos pareavam na meiose e onde não pareavam.

Herança ligada ao X

Para nosso primeiro exemplo de ligação ao X, voltaremos à cor de olho em *Drosophila*. A cor de olho tipo selvagem de *Drosophila* é vermelha, mas linhagens puras de olhos brancos estão disponíveis (Figura 2.26). Essa diferença fenotípica é determinada por dois alelos de um gene situado na região diferencial do cromossomo X. O alelo mutante no presente caso é w para olhos brancos (a letra minúscula indica que o alelo é recessivo) e o alelo tipo selvagem correspondente é w^+ . Quando os machos de olhos brancos são cruzados com fêmeas de olho vermelho, toda a prole F_1 tem olhos vermelhos, sugerindo que o alelo para olhos brancos é recessivo. O cruzamento desses machos F_1 de olhos vermelhos e fêmeas produz uma proporção F_2 de 3:1 de moscas de olhos vermelhos para moscas de olhos brancos, mas *todas as moscas de olhos brancos são machos*. Esse padrão de herança, que apresenta uma clara diferença entre os sexos, é explicado na Figura 2.27. A base do padrão de herança é que todas as moscas F_1 recebem um alelo tipo selvagem de suas mães, mas as fêmeas de F_1 recebem um alelo tipo selvagem de seus pais. Assim, todas as fêmeas de F_1 são heterozigotas tipo selvagem (w^+/w), e os machos de F_1 são hemizigotos tipo selvagem (w^+). As fêmeas de F_1 transmitem o alelo de olho branco para metade de seus filhos, que o expressam, e para metade de suas filhas, que não o expressam porque devem herdar o alelo tipo selvagem de seus pais.

O cruzamento recíproco dá um resultado diferente; isto é, o cruzamento entre fêmeas de olhos brancos e machos de olhos vermelhos dá uma F_1 na qual todas as fêmeas são de olhos vermelhos, mas todos os machos são de olhos brancos. Nesse caso, cada fêmea herda o alelo dominante w^+ do cromossomo X paterno, enquanto cada macho herda o alelo recessivo w de sua mãe. A F_2 consiste em moscas metade de olhos vermelhos e metade de olhos brancos de ambos os sexos. Assim, na ligação ao sexo, vemos exemplos não só de proporções diferentes em sexos diferentes, mas também de diferenças entre cruzamentos recíprocos.

Note que a cor de olho em *Drosophila* não tem nada a ver com a determinação do sexo, e, assim, temos uma ilustração do princípio de que os genes nos cromossomos sexuais não estão necessariamente relacionados à função sexual. O mesmo é verdadeiro em humanos: na discussão da análise de heredogramas mais adiante neste capítulo, veremos muitos genes ligados ao X, embora poucos possam ser conectados à função sexual.

O alelo anormal associado à cor de olho em *Drosophila* é recessivo, mas os alelos anormais dos genes no cromossomo X que são dominantes também surgem, tais como o mutante de *Drosophila* asa com pêlos (Hw). Em tais casos, o alelo tipo selvagem (Hw^+) é recessivo. Os alelos anormais dominantes apresentam o padrão de herança correspondente ao alelo tipo selvagem para olhos vermelhos no exemplo precedente. As proporções obtidas são as mesmas.

A herança ligada ao sexo mostra regularmente proporções fenotípicas diferentes nos dois sexos da prole, bem como proporções diferentes em cruzamentos recíprocos.



FIG. 2.26 O olho vermelho da mosca é o tipo selvagem, e o olho branco é um mutante. [Carolina Biological Supply.]

Historicamente, nas primeiras décadas do século XX, a demonstração por Morgan da herança ligada ao X de *olhos brancos* na *Drosophila* foi uma peça importante de evidência que sugere que os genes estão de fato situados nos cromossomos, porque o padrão de herança foi correlacionado a um par específico de cromossomos. A idéia tornou-se conhecida como a “teoria cromossômica da herança”. Nesse período da história, recentemente foi demonstrado que, em muitos organismos, o sexo é determinado por um cromossomo X e um Y, e que, nos machos, esses cromossomos segregam igualmente na meiose para regenerar números iguais de machos e fêmeas na geração seguinte. Morgan reconheceu que a herança dos alelos do gene de cor de olho é exatamente paralelo a herança dos cromossomos X na meiose: assim, o gene estava provavelmente no cromossomo X. A herança dos *olhos brancos* foi estendida para as linhagens de *Drosophila* que tinham números anormais de cromossomos sexuais. Com o uso dessa nova situação, será possível prever os padrões de herança de genes a partir da segregação de cromossomos anormais. Que essas previsões estavam corretas foi um teste convincente da teoria cromossômica.

Outras análises genéticas revelaram que, em galinhas e mariposas, a herança ligada ao sexo podia ser explicada apenas se a fêmea fosse o sexo heterogamético. Nesses organismos, os cromossomos sexuais femininos são chamados de ZW e os masculinos foram chamados de ZZ.

► 2.6 Análise de Heredogramas

As reproduções humanas, como a dos organismos experimentais, fornecem muitos exemplos de herança monogênica. Entretanto, os cruzamentos experimentais controlados não podem ser feitos em humanos, e, assim, os geneticistas têm de recorrer aos registros médicos na esperança de que as reproduções informativas que tenham ocorrido (tais como reproduções monoíbridas) possam ser usadas para deduzir herança monogênica. Tal levantamento dos registros de reproduções é chamado de **análise de heredogramas**. Um membro de uma família que primeiro chama a atenção de

Um exemplo de herança ligada ao X

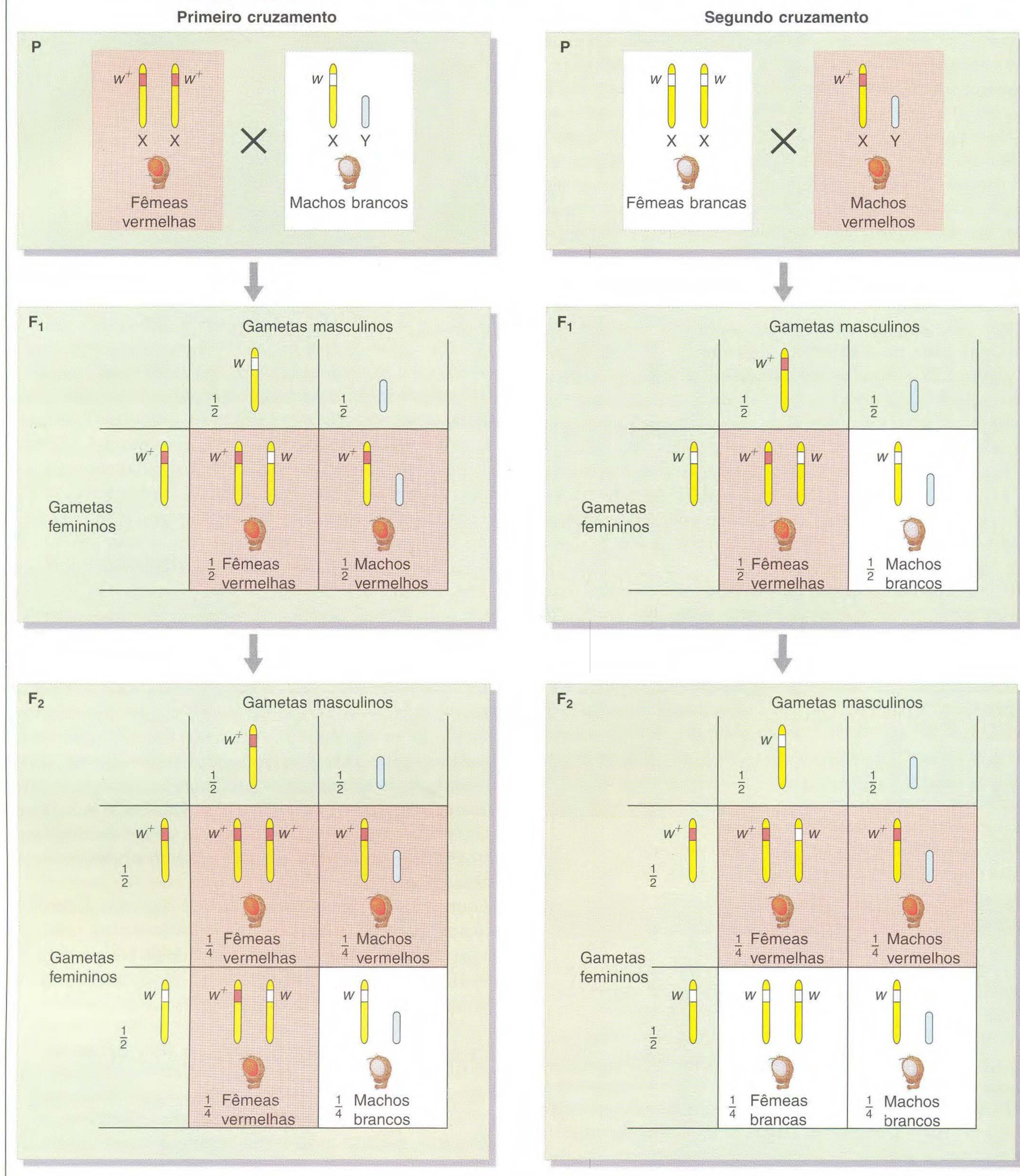


FIG. 2.27 Cruzamentos recíprocos entre *Drosophila* de olho vermelho (vermelho) e olho branco (branco) dão resultados diferentes. Os alelos são ligados ao X, e a herança do cromossomo X explica as proporções fenotípicas observadas, que são diferentes daquelas dos genes autossômicos. (Em *Drosophila* e muitos outros sistemas experimentais, um expoente com sinal + é usado para designar o alelo normal, ou tipo selvagem. Aqui, w^+ codifica olhos vermelhos e w codifica olhos brancos.)



FIG. 2.28 Uma variedade de símbolos são usados nas análises de heredogramas. [De W.F. Bodmer and L.L. Cavalli-Sforza, *Genetics, Evolution, and Man*. Copyright 1976 by W.H. Freeman and Company.]

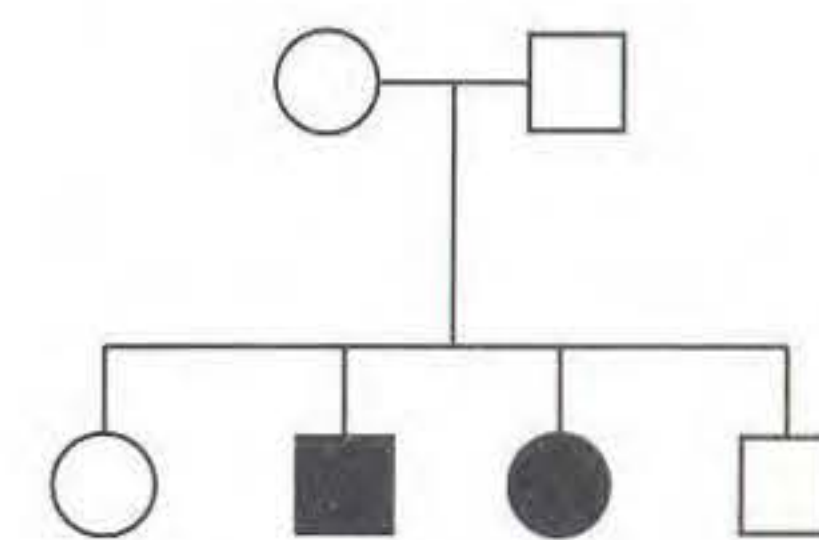
um geneticista é chamado de **propósito**. Geralmente, o fenótipo do propósito é excepcional de algum modo. Por exemplo, o propósito pode sofrer de algum tipo de distúrbio médico. O pesquisador então traça a história do fenótipo através da história da família e desenha uma árvore genealógica, ou heredograma, usando os símbolos padrão dados na Figura 2.28.

Para ver a herança monogênica, os padrões no heredograma têm que ser interpretados de acordo com a lei de Mendel da segregação igual, mas os humanos geralmente têm poucos filhos, e, assim, devido a essa prole pequena, as proporções esperadas de 3:1 e 1:1 geralmente não são vistas a menos que muitos heredogramas similares sejam combinados. O enfoque à análise de heredogramas também depende de se um dos fenótipos contrastantes é um distúrbio raro ou ambos os fenótipos de um par são morfos comuns de um polimorfismo. A maioria dos heredogramas são feitos por motivos médicos, e, daí, o interesse por distúrbios médicos que são, quase por definição, raros. Nesse caso, temos dois fenótipos: a presença e a ausência do distúrbio. Quatro padrões de herança monogênica são relevantes nos heredogramas. Vejamos, primeiro, os distúrbios recessivos causados pelos alelos recessivos de genes autossômicos únicos.

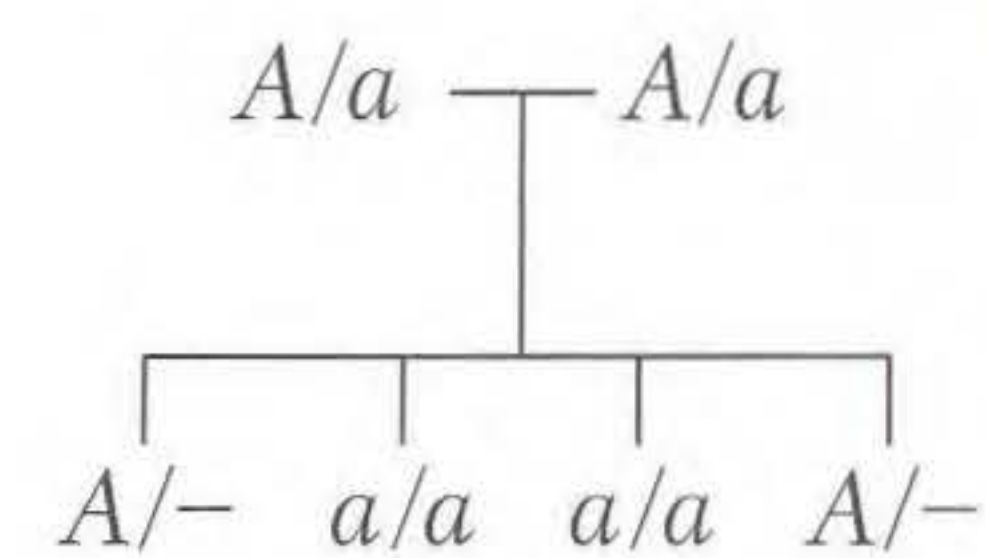
Distúrbios autossômicos recessivos

O fenótipo afetado de um distúrbio autossômico recessivo é herdado como um alelo recessivo. Assim, o fenótipo correspondente não-afetado deve ser herdado como o alelo dominante correspondente. Por exemplo, a doença humana fenilcetonúria, já discutida, é herdada de modo mendeliano simples como um fenótipo recessivo, com a PKU determinada pelo alelo p e a condição normal determinada por P . Portanto, os que sofrem dessa doença são de fenótipo p/p , e as pessoas que não têm a doença são ou P/P ou P/p . Note que o termo tipo selvagem e seus símbolos alélicos não são usados em genética humana, pois o tipo selvagem é impossível de definir.

Que padrões em um heredograma revelariam uma herança autossômica recessiva? Os dois pontos-chave são que (1) geralmente o distúrbio aparece na prole de genitores não-afetados, e (2) a prole afetada inclui tanto homens quanto mulheres. Quando sabemos que tanto homens quanto mulheres são afetados, podemos deduzir que, mais provavelmente, estamos lidando com uma herança mendeliana simples de um gene em um autossomo, e não de um gene em um cromossomo sexual. O seguinte heredograma típico ilustra o ponto-chave de que crianças afetadas nascem de genitores não-afetados:



Por esse padrão, podemos deduzir uma reprodução monóibrida simples, com o alelo recessivo responsável pelo fenótipo excepcional (indicado em preto). Ambos os genitores devem ser heterozigotos, digamos, A/a ; ambos devem ter um alelo a , pois cada um contribui com um alelo a para cada criança afetada, e ambos devem ter um alelo A porque são fenotipicamente normais. Podemos identificar os genótipos das crianças (na ordem mostrada) como $A/-$, a/a , a/a e $A/-$. Assim, o heredograma pode ser reescrito como:



Esse heredograma não apóia a hipótese de herança recessiva ligada ao X, porque, sob essa hipótese, uma filha afetada teria que ter uma mãe heterozigota (possível) e um pai hemizigoto, o que é claramente impossível porque o pai teria que ter expressado o fenótipo do distúrbio.

Note que, mesmo as regras mendelianas estando em funcionamento, as proporções mendelianas não são necessariamente observadas em famílias isoladas, devido ao pequeno tamanho da amostra, como previsto antes. No capítulo precedente, observamos uma proporção fenotípica de 1:1 na prole de uma reprodução monóibrida. Se o casal tivesse, digamos, 20 filhos, a proporção seria algo

como 15 filhos não-afetados e 5 com PKU (uma proporção de 3:1), mas, em uma pequena amostra de 4 filhos, qualquer proporção é possível, e todas as proporções são comumente encontradas.

Os heredogramas de distúrbios autossômicos recessivos tendem a ser pequenos, com poucos símbolos escuros. Uma condição recessiva apresenta-se em grupos de irmãos afetados, e as pessoas nas gerações anteriores e posteriores tendem a não ser afetadas. Para compreender porque isso é assim, é importante ter alguma compreensão da estrutura genética das populações subjacentes a essas condições raras. Por definição, se a condição é rara, a maioria das pessoas não possuem o alelo anormal. Além disso, a maioria das pessoas que possuem o alelo anormal são heterozigotas para ele, e não homozigotas. O motivo básico de os heterozigotos serem muito mais comuns que os homozigotos recessivos é que, para ser homozigoto recessivo, ambos os genitores têm que ter um alelo a , mas, para ser um heterozigoto, apenas um genitor deve tê-lo.

O nascimento de uma pessoa afetada geralmente depende da rara união aleatória de genitores heterozigotos não-aparentados. Entretanto, a endogamia (reprodução entre parentes) aumenta a chance de que dois heterozigotos se reproduzam. Um exemplo de um casamento entre primos é mostrado na Figura 2.29. As pessoas III-5 e III-6 são primos em primeiro grau e produzem dois homozigotos para o alelo raro. Você pode ver, na Figura 2.29, que um ancestral que é heterozigoto pode produzir muitos descendentes que também são

heterozigotos. Assim, dois primos podem ter o mesmo alelo recessivo herdado de um ancestral comum. Para que duas pessoas não-aparentadas sejam heterozigotas, elas teriam que herdar o alelo raro de ambas as suas famílias. Assim, a reprodução entre parentes geralmente corre um risco maior de produzir distúrbios recessivos do que as reproduções entre não-parentes. Por esse motivo, os casamentos de primos em primeiro grau contribuem com uma grande proporção dos que sofrem de doenças recessivas na população.

Quais são alguns exemplos de distúrbios recessivos humanos? A fibrose cística é uma doença herdada de acordo com as regras mendelianas como um fenótipo autossômico recessivo. Seu sintoma mais importante é a secreção de grandes quantidades de muco nos pulmões, resultando em morte por uma combinação de efeitos, mas geralmente precipitada por infecção das vias respiratórias. O muco pode ser deslocado por batidas mecânicas no tórax, e a infecção pulmonar pode ser evitada por antibióticos. Assim, com tratamento, os pacientes com fibrose cística podem viver até a vida adulta. O gene da fibrose cística foi um dos primeiros genes de doença humana a serem isolados ao nível do DNA, em 1989. Essa linha de pesquisa revelou que o distúrbio é causado por uma proteína defeituosa que transporta íons cloreto através da membrana celular. A alteração resultante do balanço de sal muda a constituição do muco do pulmão. Essa nova compreensão do funcionamento do gene em pessoas afetadas e não-afetadas deu esperanças para um tratamento mais efetivo.

O albinismo humano também é herdado do modo autossômico recessivo padrão. O alelo mutante é de um gene que, normalmente, sintetiza o pigmento melanina marrom ou preto. O padrão de herança desse gene haplossuficiente é mostrado juntamente com o defeito celular na Figura 1.8.

Homozigotos recessivos de endogamia

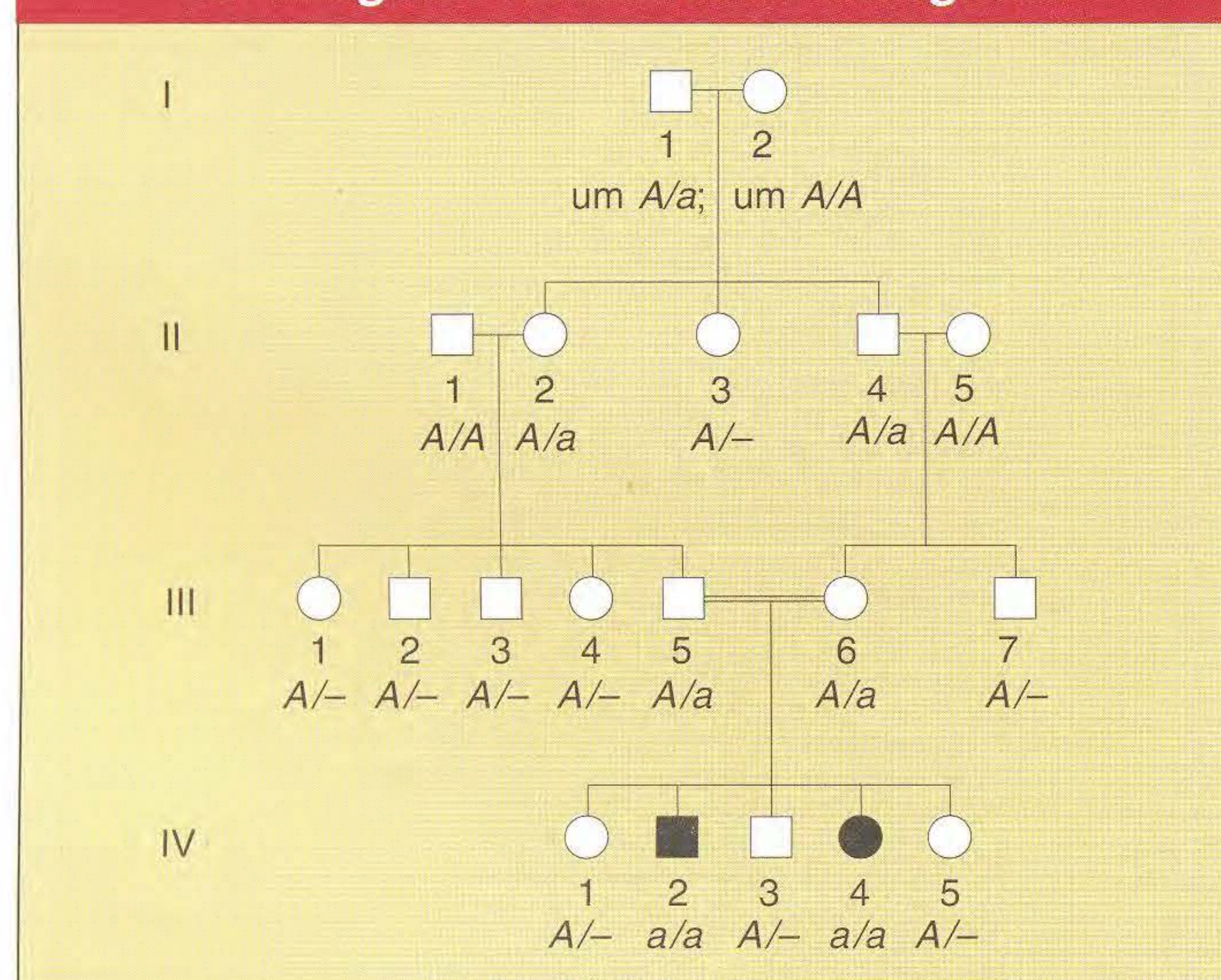


FIG. 2.29 Heredograma de um fenótipo recessivo raro determinado por um alelo recessivo a . Os símbolos genéticos normalmente não são incluídos nos heredogramas, mas os genótipos são inseridos aqui para referência. As pessoas II-1 e II-5 entraram para a família; elas são supostas como sendo normais porque a condição herdável examinada é rara. Note também que não é possível avaliar o genótipo; algumas pessoas são indicadas por $A/-$. As pessoas III-5 e III-6, que geraram os recessivos na geração IV, são primos em primeiro grau. Ambas obtiveram seu alelo recessivo de um avô, seja I-1 ou I-2.

Em heredogramas, um distúrbio autossômico recessivo é revelado pelo surgimento do distúrbio na prole masculina e feminina de genitores não-afetados.

Distúrbios autossômicos dominantes

Que padrões de heredogramas são esperados de distúrbios autossômicos dominantes? Aqui, o alelo normal é recessivo, e o alelo defeituoso é dominante. Pode ser paradoxal que um distúrbio raro possa ser dominante, mas lembre que a dominância e a recessividade são simples propriedades de como os alelos atuam nos heterozigotos e não são definidos em relação ao quanto eles são comuns na população. Um bom exemplo de um fenótipo dominante raro que apresenta herança monogênica é a pseudo-acondroplasia, um tipo de nanismo (Figura 2.30). Com relação a esse gene, as pessoas com estatura normal são genotipicamente d/d , e o fenótipo nanismo em princípio pode ser D/d ou D/D . Entretanto, as duas “doses” do alelo D no genótipo D/D são tidas como produzindo um efeito tão grave que esse genótipo é letal. Se essa crença é geralmente verdadeira, todos os indivíduos com nanismo são heterozigotos.

Na análise de heredogramas, os principais indícios para identificação de distúrbios autossômicos dominantes com

Fenótipo de pseudo-acondroplasia



FIG. 2.30 O fenótipo de pseudo-acondroplasia humana é ilustrado aqui por uma família de cinco irmãs e dois irmãos. O fenótipo é determinado por um alelo dominante, que podemos chamar de *D*, o qual interfere com o crescimento dos ossos longos durante o desenvolvimento. Essa foto foi tirada quando a família chegou em Israel após o final da Segunda Guerra Mundial. [UPI/Bettmann News Photos.]

herança mendeliana são que o fenótipo tende a aparecer em cada geração do heredograma e que os pais e mães afetados transmitem o fenótipo para os filhos e filhas. Novamente, a representação igual de ambos os sexos entre a prole afetada exclui a herança pelos cromossomos sexuais. O fenótipo aparece em cada geração porque, em geral, o alelo anormal levado por uma pessoa deve ter vindo de um genitor na geração anterior. (Alelos anormais também podem surgir *de novo* por mutações. Deve-se ter em mente essa possibilidade para distúrbios que interferem na reprodução porque, aqui, a condição improvavelmente foi herdada de um genitor afetado.) Um heredograma típico de um distúrbio dominante é mostrado na Figura 2.31. Novamente, note que as proporções mendelianas não são necessariamente observadas nas famílias. Como nos distúrbios recessivos, as pessoas portadoras de uma cópia de um alelo raro *A* (*A/a*) são muito mais comuns que os portadores de duas cópias (*A/A*); logo, a maioria das pessoas afetadas são heterozigotas, e praticamente todas as reproduções que geram prole com distúrbios dominantes são *A/a* × *a/a*. Portanto, se a prole de tais casamentos forem somadas, é esperada uma proporção de 1:1 de pessoas não afetadas (*a/a*) para afetadas (*A/a*).

A doença de Huntington é um exemplo de uma doença herdada como um fenótipo dominante determinado por um único alelo. O fenótipo é de degeneração neural, levando a convulsões e morte prematura. O cantor popular Woody Guthrie sofria de doença de Huntington. A doença é incomum por se manifestar tardiamente; os sintomas geralmente

só aparecem depois que a pessoa já teve filhos (Figura 2.32). Quando a doença foi diagnosticada em um genitor, cada filho já nascido sabe que ele ou ela tem uma chance de 50% de herdar o alelo e a doença associada. Esse padrão trágico inspirou um grande esforço em achar meios de identificar pessoas que possuem o alelo anormal antes que manifestem o início da doença. Hoje existem diagnósticos moleculares para identificar pessoas que possuem o alelo Huntington.

Herança de um distúrbio autossômico dominante

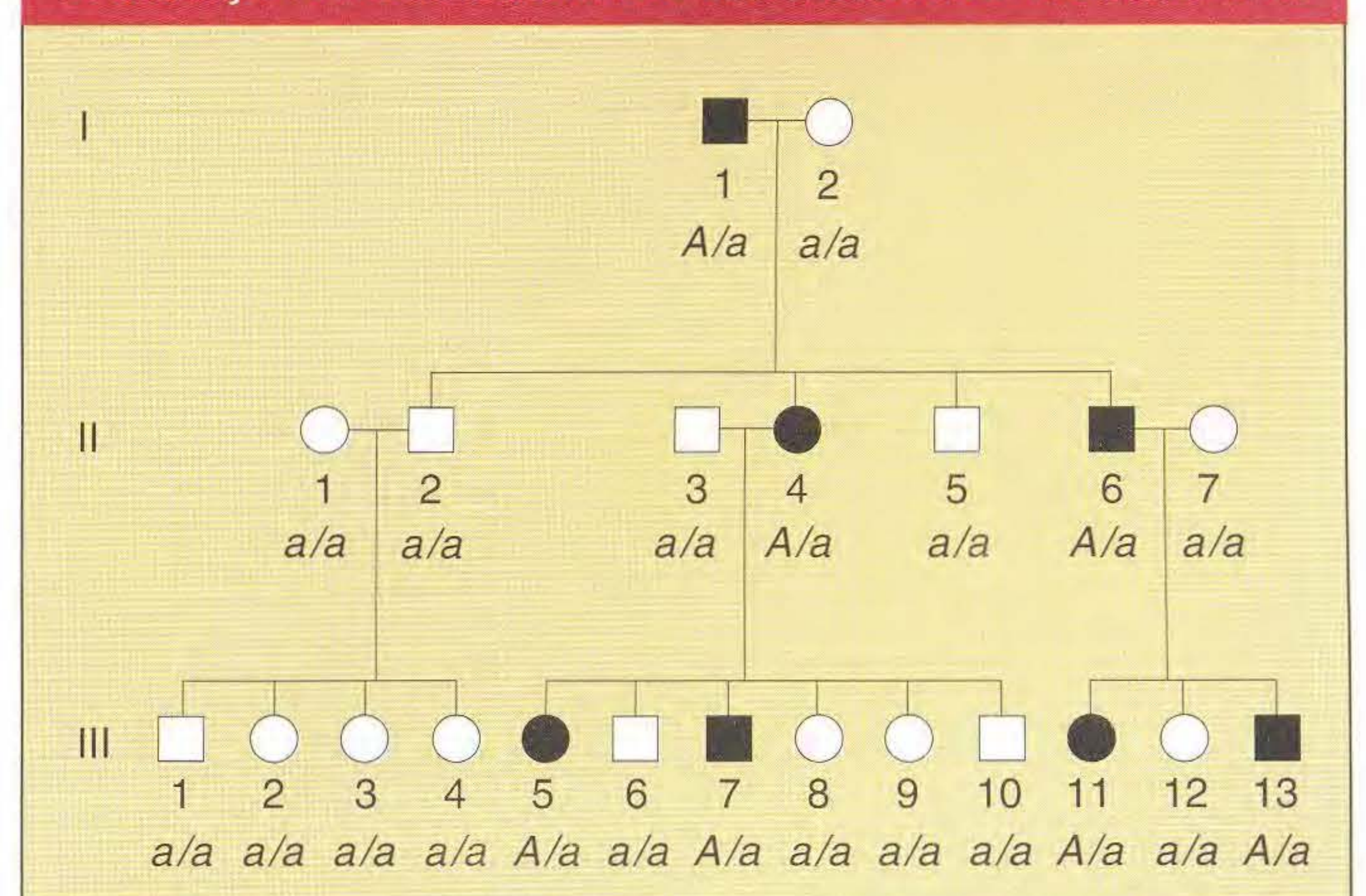


FIG. 2.31 Heredograma de um fenótipo dominante determinado por um alelo dominante *A*. Nesse heredograma, todos os genótipos foram deduzidos.

Doença de Huntington de manifestação tardia

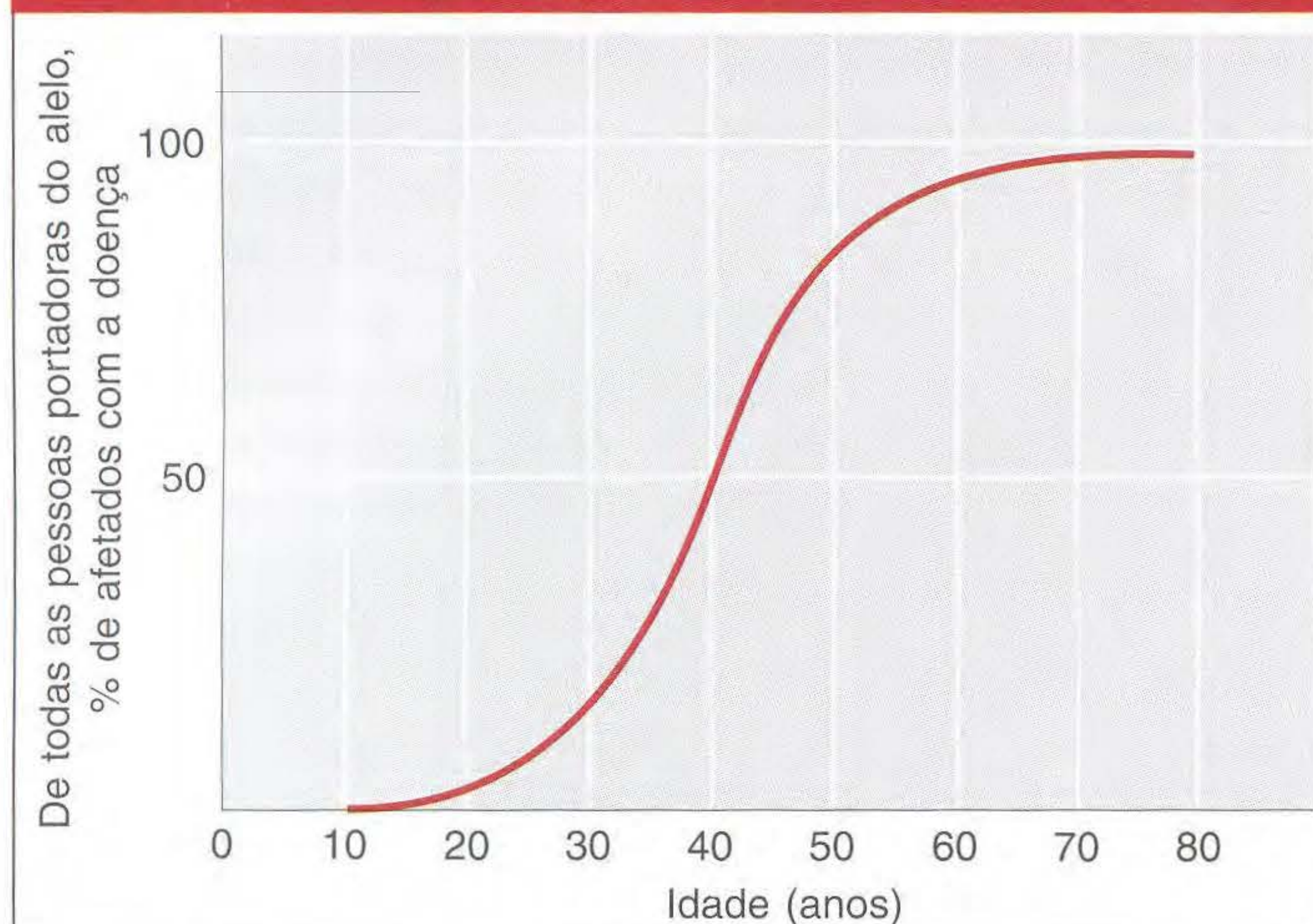


FIG. 2.32 O gráfico mostra que as pessoas portando o alelo em geral não manifestam a doença até após a idade de criar filhos.

Algumas outras condições dominantes raras são a polidactilia (dedos extras), mostrada na Figura 2.33, e o piebaldismo, mostrado na Figura 2.34.

Os heredogramas de distúrbios mendelianos autossômicos dominantes mostram homens e mulheres afetados em cada geração; também mostram homens e mulheres afetados transmitindo a condição a proporções iguais de seus filhos e filhas.

Polimorfismos autossômicos

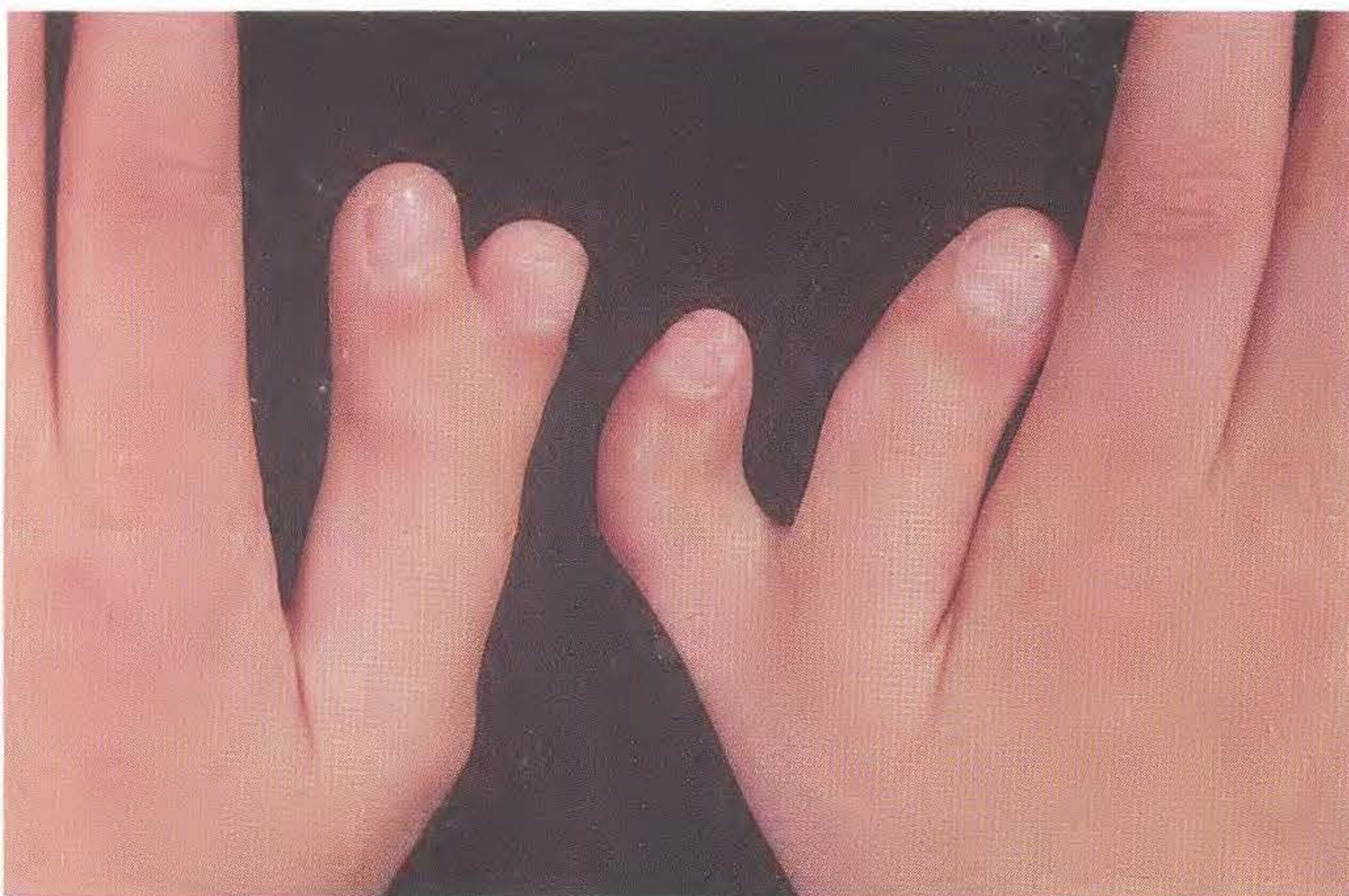
Em populações naturais de organismos, um polimorfismo é a coexistência de dois ou mais fenótipos comuns de uma

característica. Os fenótipos alternativos de um **polimorfismo (morfo)** são geralmente herdados como alelos de um único gene autossômico no modo mendeliano padrão. Entre os muitos exemplos humanos estão os seguintes dimorfismos: olhos azuis \times castanhos, cabelo pigmentado \times loiro, covinhas na bochecha \times nenhuma, bico de viúva \times nenhum, lobos da orelha livres \times presos. Em cada exemplo, o morfo é determinado pelo alelo dominante escrito primeiro.

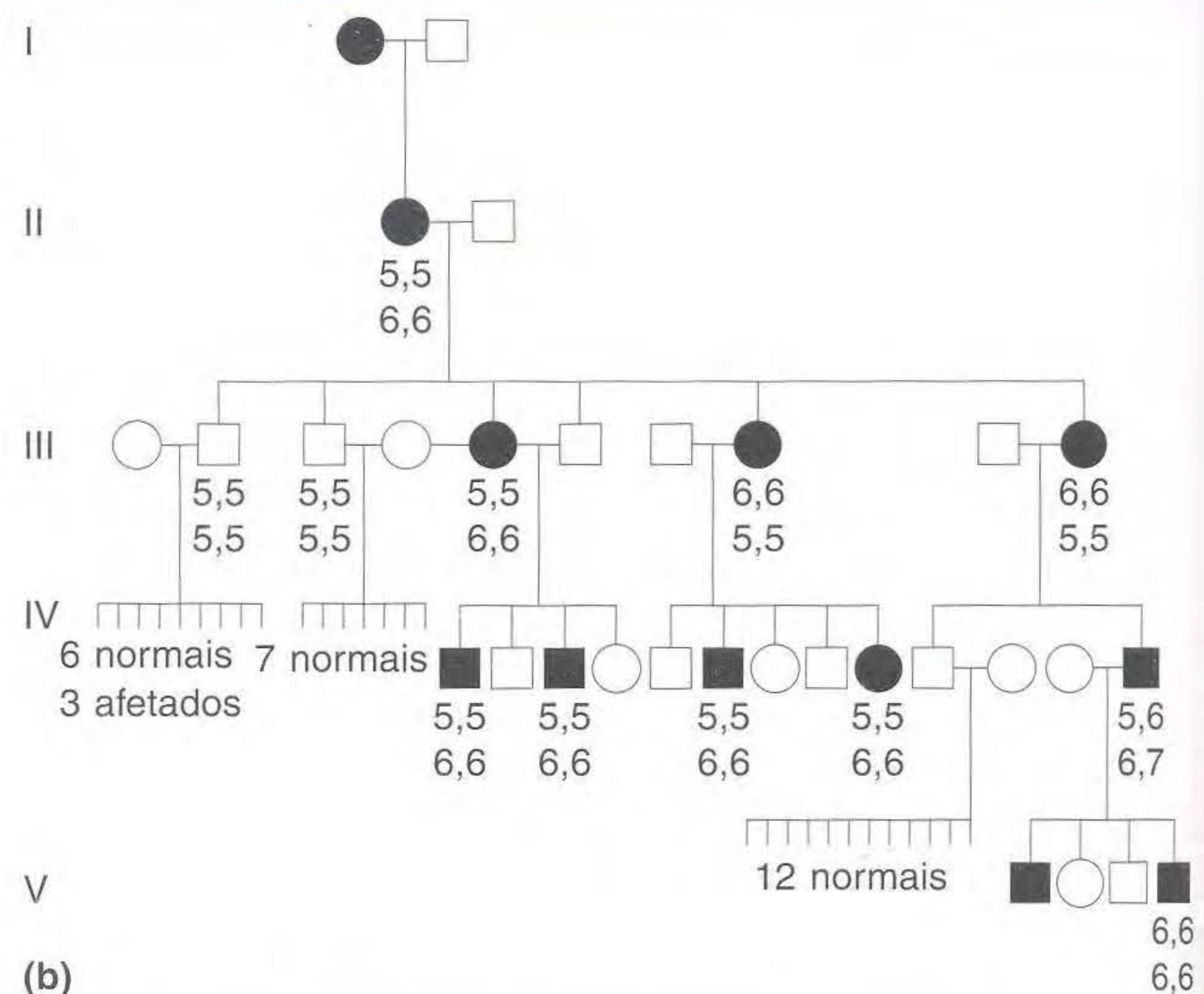
A interpretação dos heredogramas para polimorfismos é um pouco diferente do de distúrbios raros porque, por definição, os morfismos são comuns. Vejamos um heredograma para um interessante caso humano. Um **dimorfismo** é o tipo mais simples de polimorfismo, apenas com dois morfismos. A maioria das populações humanas são dimórficas para a habilidade de sentir o gosto da feniltiocarbamida (PTC); isto é, as pessoas podem detectá-la como tendo um gosto amargo, ou, para a grande surpresa e descrença dos que sentem, não sentir nada. Pelo heredograma na Figura 2.35, podemos ver que dois sensíveis às vezes têm filhos insensíveis, o que torna claro que o alelo que confere a habilidade de sentir é dominante, e que o alelo de não sentir é recessivo. Note, na Figura 2.35, que quase todas as pessoas que entraram para a família têm o alelo recessivo, seja na condição heterozigota ou homozigota. Tal heredograma, portanto, difere dos distúrbios recessivos raros, para os quais a suposição convencional é que todos os que entram para a família são homozigotos normais. Como ambos os alelos de PTC são comuns, não surpreende que todos menos um membro da família casou-se com pessoas com pelo menos uma cópia do alelo recessivo.

O polimorfismo é um fenômeno genético interessante. Os geneticistas de populações ficaram surpresos com tanto polimorfismo que existe nas populações naturais de plantas e animais em geral. Além disso, muito embora a genética de

Polidactilia



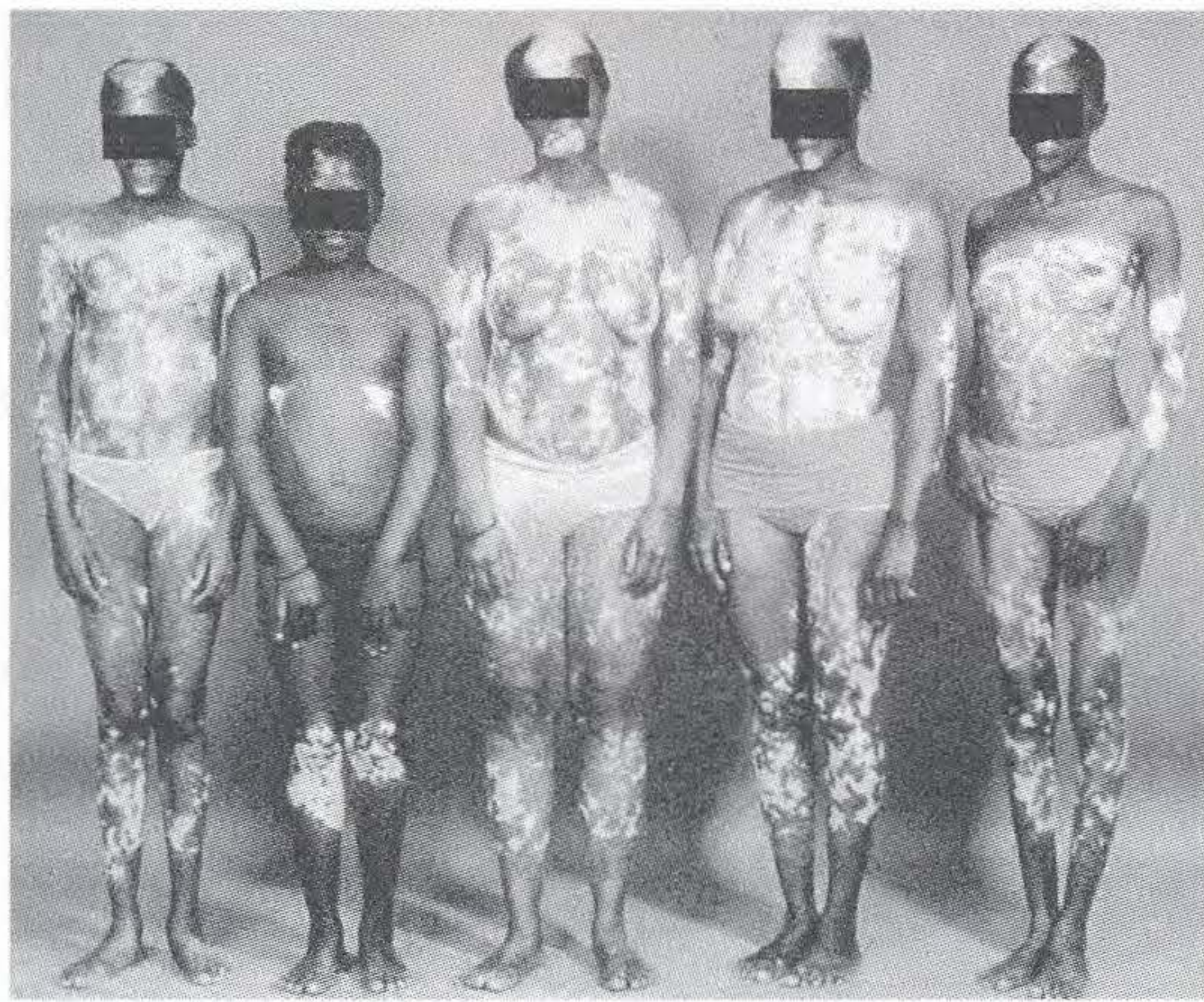
(a)



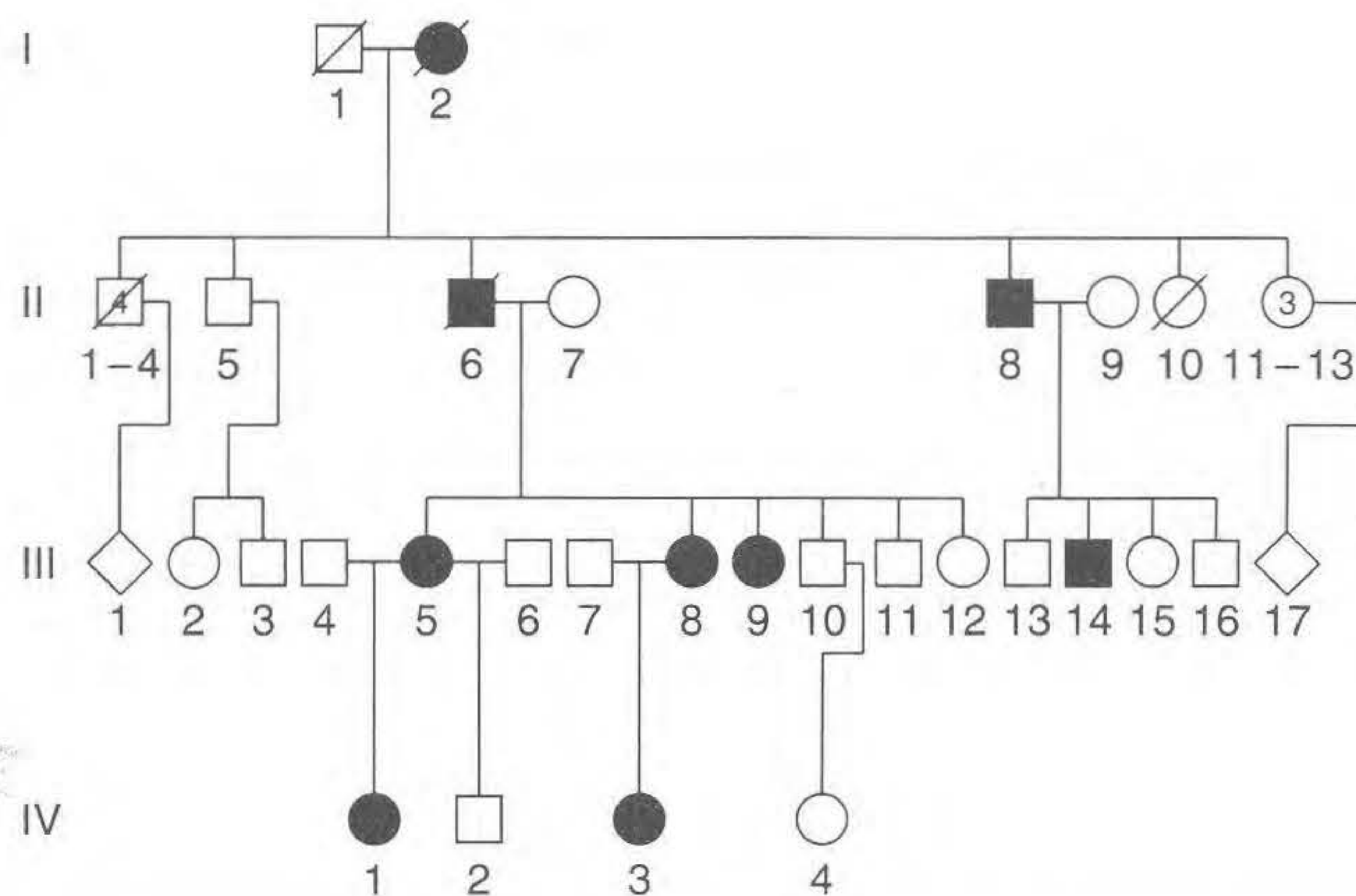
(b)

FIG. 2.33 A polidactilia é um fenótipo raro dominante das mãos e pés humanos. (a) Polidactilia, caracterizada por dedos e artelhos extras, ou ambos, é determinada por um alelo *P*. Os números no heredograma (b) dão o número de dedos nas linhas superiores e o número de artelhos nas inferiores. (Note a variação na expressão do alelo *P*.) [(a) Foto de © Biophoto Associates/Science Source.]

Manchas de piebaldismo dominante



(a)



(b)

FIG. 2.34 As manchas do piebaldismo são um fenótipo humano dominante raro. Embora o fenótipo seja encontrado esporadicamente em todas as raças, os padrões mostram-se mais nos indivíduos de pele escura. (a) A foto mostra as imagens de frente e de costas das pessoas afetadas IV-1, IV-3, III-5, III-8 e III-9 do (b) heredograma familiar. Note a variação na expressão do gene de piebaldismo entre os membros da família. Os padrões são tidos como causados pelo alelo dominante que interfere na migração de melanócitos (células produtoras de melanina) da superfície dorsal para a ventral no curso do desenvolvimento. A testa branca é particularmente característica, sendo geralmente acompanhada de uma região branca no cabelo.

O piebaldismo não é uma forma de albinismo; as células nas manchas claras têm o potencial genético de produzir melanina, mas, como não são melanócitos, elas não são desenvolvimentalmente programadas para fazê-lo. No verdadeiro albinismo, as células não têm o potencial de produzir melanina. (O piebaldismo é causado por mutação em *c-kit*, um tipo de gene chamado *proto-oncogene* que será discutido no Capítulo 15.) [(a e b) De I. Winship, K. Young, R. Martell, R. Ramesar, D. Curtis, and P. Beighton. "Piebaldism: An Autonomous Autosomal Dominant Entity". *Clin. Genet.* 39, 1991, 330.]

polimorfismos seja direta, existem muito poucos polimorfismos para os quais existam explicações satisfatórias para a coexistência dos morfos. Mas o polimorfismo é viçoso em todos os níveis da análise genética, mesmo a nível de DNA. De fato, os polimorfismos observados a nível de DNA têm sido valiosos como marcos para ajudar os geneticistas a encontrar seu caminho nos cromossomos de organismos complexos, como será descrito no Capítulo 4. A genética de populações e evolutiva de polimorfismos é considerada nos Capítulos 17 e 19.

As populações de plantas e animais (incluindo humanos) são altamente polimórficas. Morfos contrastantes são freqüentemente herdados como alelos de um único gene.

Distúrbios recessivos ligados ao X

Vejamos os heredogramas de distúrbios causados por alelos recessivos de genes situados no cromossomo X. Tais heredogramas tipicamente mostram as seguintes características:

1. Muito mais homens que mulheres apresentam o fenótipo raro em estudo. O motivo é que uma mulher herda o genótipo apenas se tanto sua mãe quanto seu pai tiverem o alelo (por exemplo, $X^A X^a \times X^a Y$), enquanto um homem pode herdar o fenótipo apenas quando a mãe possui o alelo ($X^A X^a \times X^A Y$). Se o alelo recessivo é muito raro, quase todas as pessoas que apresentam o fenótipo são homens.
2. Ninguém na prole de um homem afetado apresenta o fenótipo, mas todas as suas filhas são "portadoras", por-

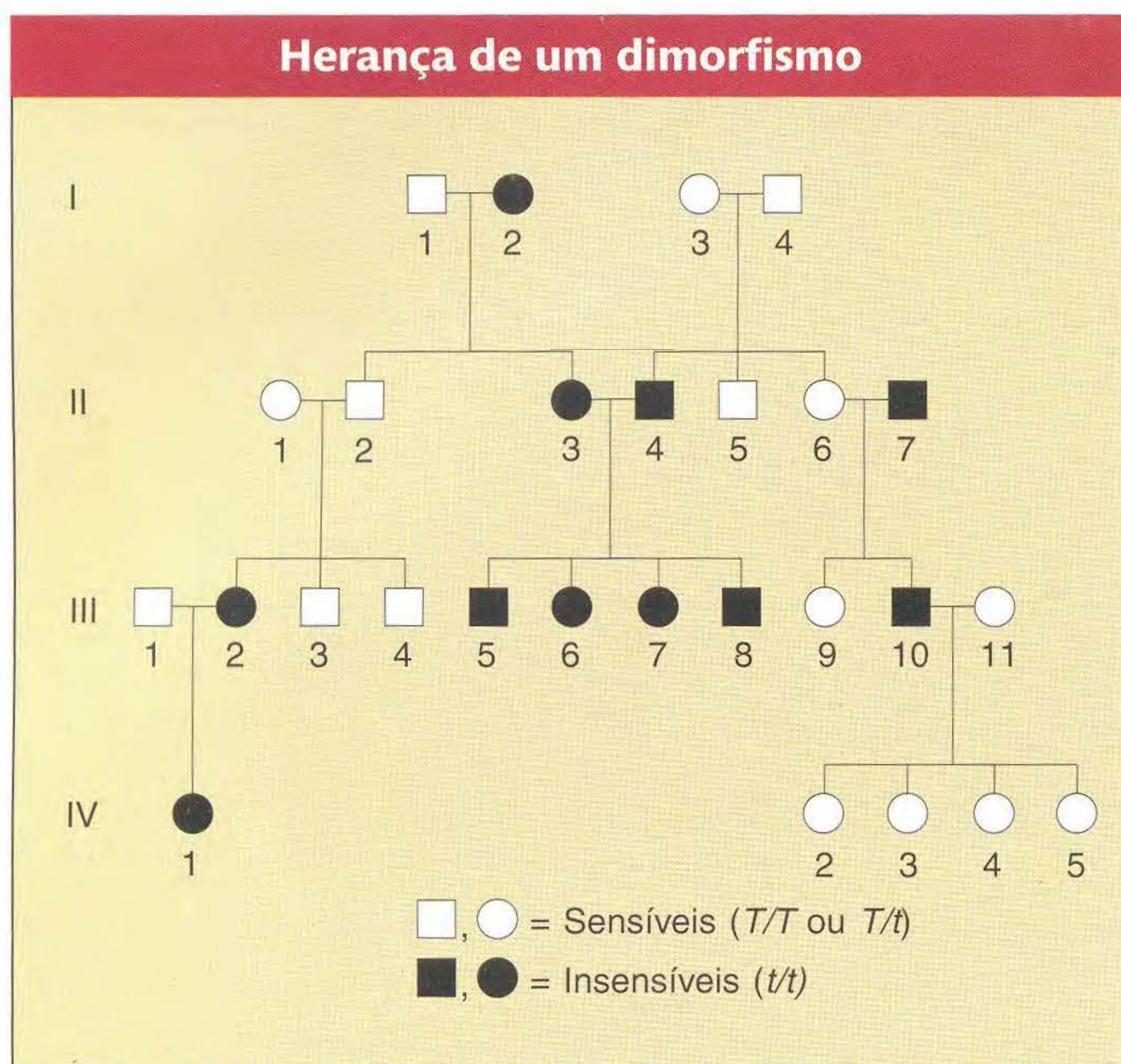


FIG. 2.35 Heredograma para a habilidade em sentir o gosto da feniltiocarbamida.

tando o alelo recessivo escondido na condição de heterozigoto. Na geração seguinte, metade dos filhos dessas filhas portadoras apresentam o fenótipo (Figura 2.36).

3. Nenhum dos filhos de um homem afetado apresenta o fenótipo em estudo, nem transmite a condição para seus descendentes. O motivo subjacente à falta de transmissão de homem para homem é que um filho obtém seu cromossomo Y de seu pai; logo, ele não pode herdar normalmente o cromossomo X do pai também. Contrariamente, a transmissão de homem para homem de um distúrbio é um diagnóstico útil para uma condição herdada autossomicamente.

Na análise de heredogramas de raros recessivos ligados ao X, uma mulher normal de genótipo desconhecido supostamente é homozigota a menos que existam evidências contrárias.

Talvez o exemplo mais familiar de herança recessiva ligada ao X é o daltonismo. As pessoas com essa condição são incapazes de distinguir o vermelho do verde. Os genes para visão em cores foram caracterizados a nível molecular. A visão em cores é baseada em três tipos diferentes de células-cone na retina, cada uma sensível aos comprimentos de onda vermelho, verde ou azul. Os determinantes genéticos para as células-cone vermelha e verde estão no cromossomo X. As pessoas com daltonismo verde e vermelho têm uma mutação em um desses dois genes. Como em qualquer distúrbio ligado ao X, existem muito mais homens com o fenótipo do que mulheres.

Outro exemplo familiar é a *hemofilia*, a falha de coagulação do sangue. Muitas proteínas atuam em sequência para fazer o sangue coagular. O tipo mais comum de hemofilia é causado pela ausência ou mal funcionamento de uma dessas proteínas de coagulação, chamada *fator VIII*. Um heredograma bem conhecido de hemofilia é o das famílias reais da

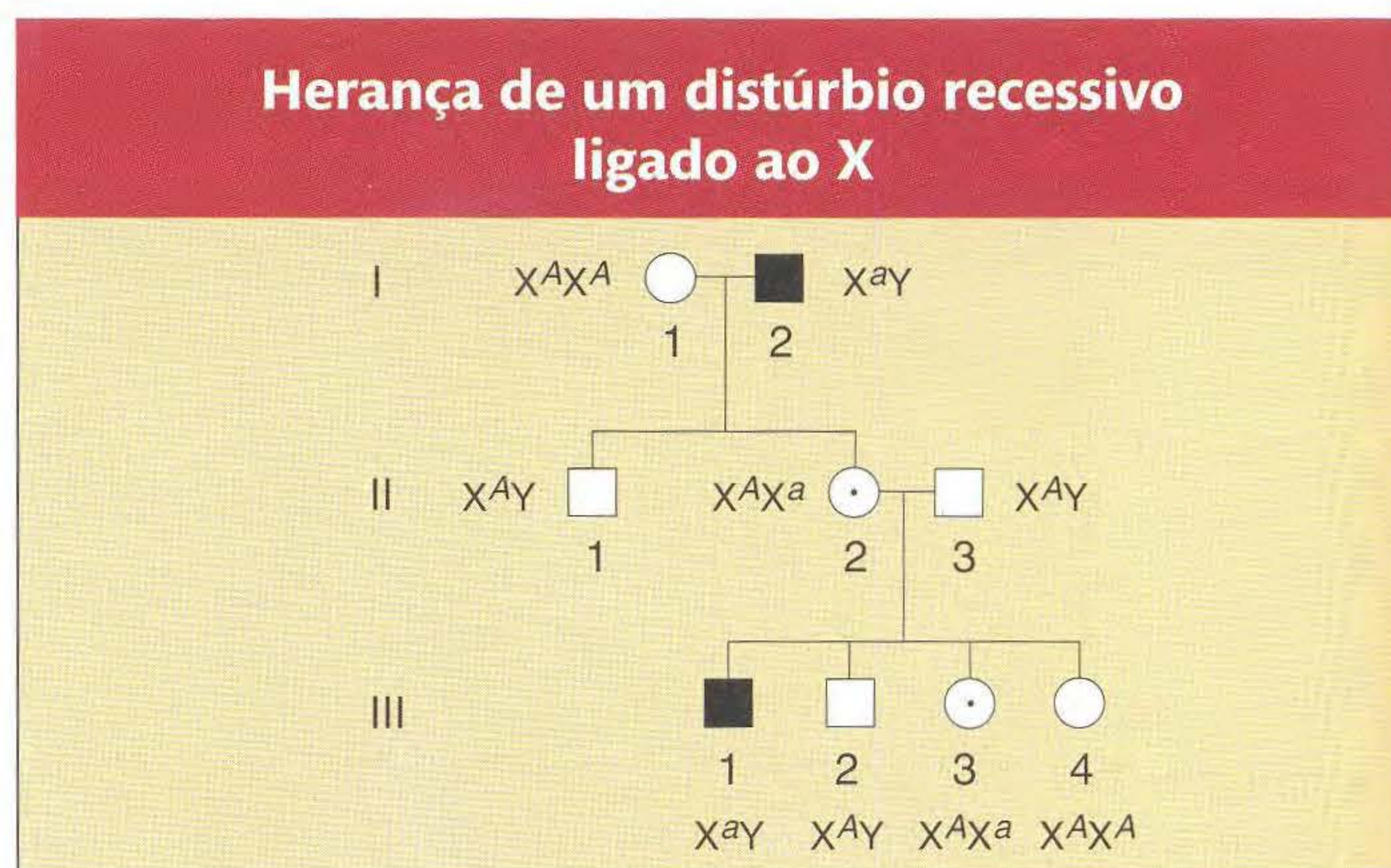


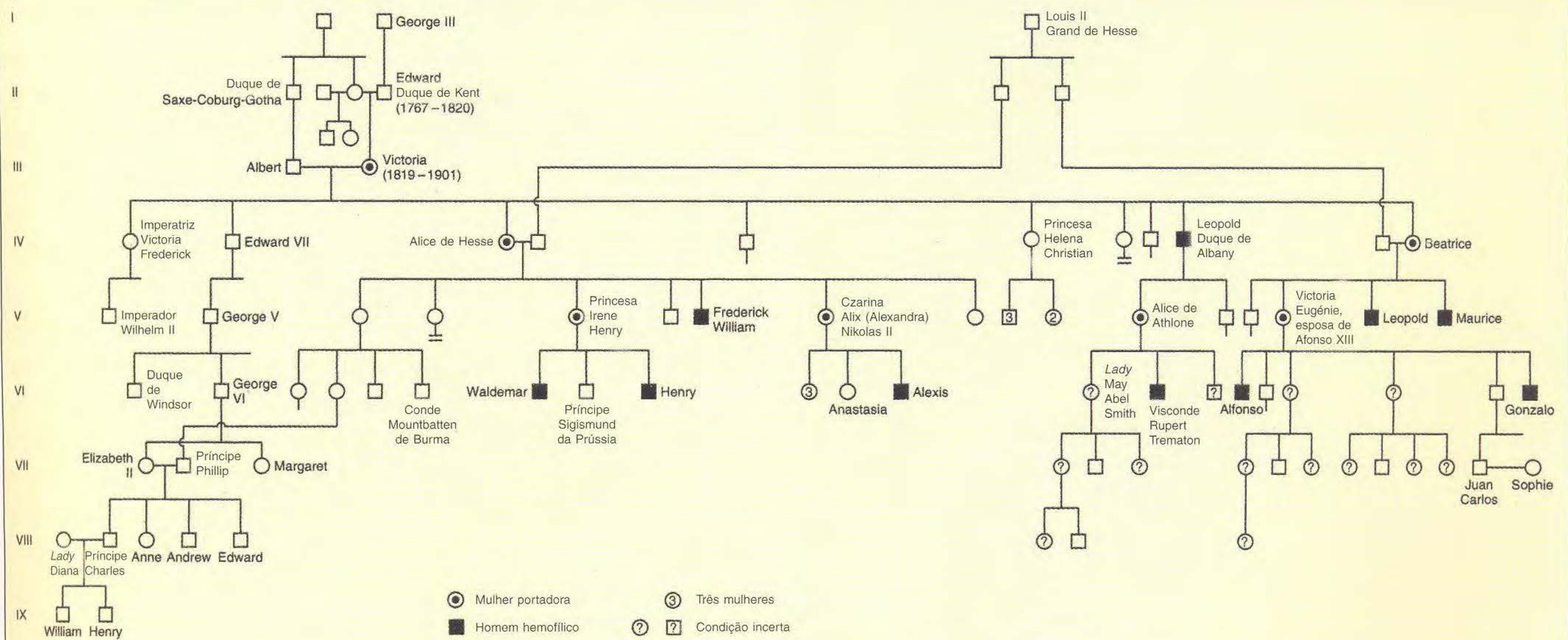
FIG. 2.36 Como geralmente é o caso, a expressão de alelos recessivos ligados ao X ocorre apenas nos homens. Esses alelos não são expressos pelas filhas na geração seguinte, para serem novamente expressos nos filhos. Note que III-3 e III-4 não podem ser distintos fenotipicamente.

Europa (Figura 2.37). O alelo original da hemofilia no heredograma surgiu espontaneamente como uma mutação nas células reprodutivas ou em um dos pais da rainha Victoria ou na própria rainha Victoria. Entretanto, alguns propuseram que a origem do alelo foi um amante secreto da mãe de Victoria. Alexis, o filho do czar da Rússia, herdou o alelo da hemofilia da rainha Victoria, que era avó de sua mãe, Alexandra. Hoje em dia, a hemofilia pode ser tratada medicamente, mas antes era uma condição potencialmente fatal. É interessante notar que o Talmud judaico contém regras sobre as exceções da circuncisão masculina, mostrando claramente que o modo de transmissão da doença por mulheres portadoras não-afetadas era bem compreendido naquelas épocas antigas. Por exemplo, uma exceção era para os filhos de mulheres cujos filhos de irmãs sangravam profundamente quando submetidos à circuncisão. Assim, o sangramento anormal era conhecido como sendo transmitido pelas mulheres da família, mas expresso apenas em seus filhos homens.

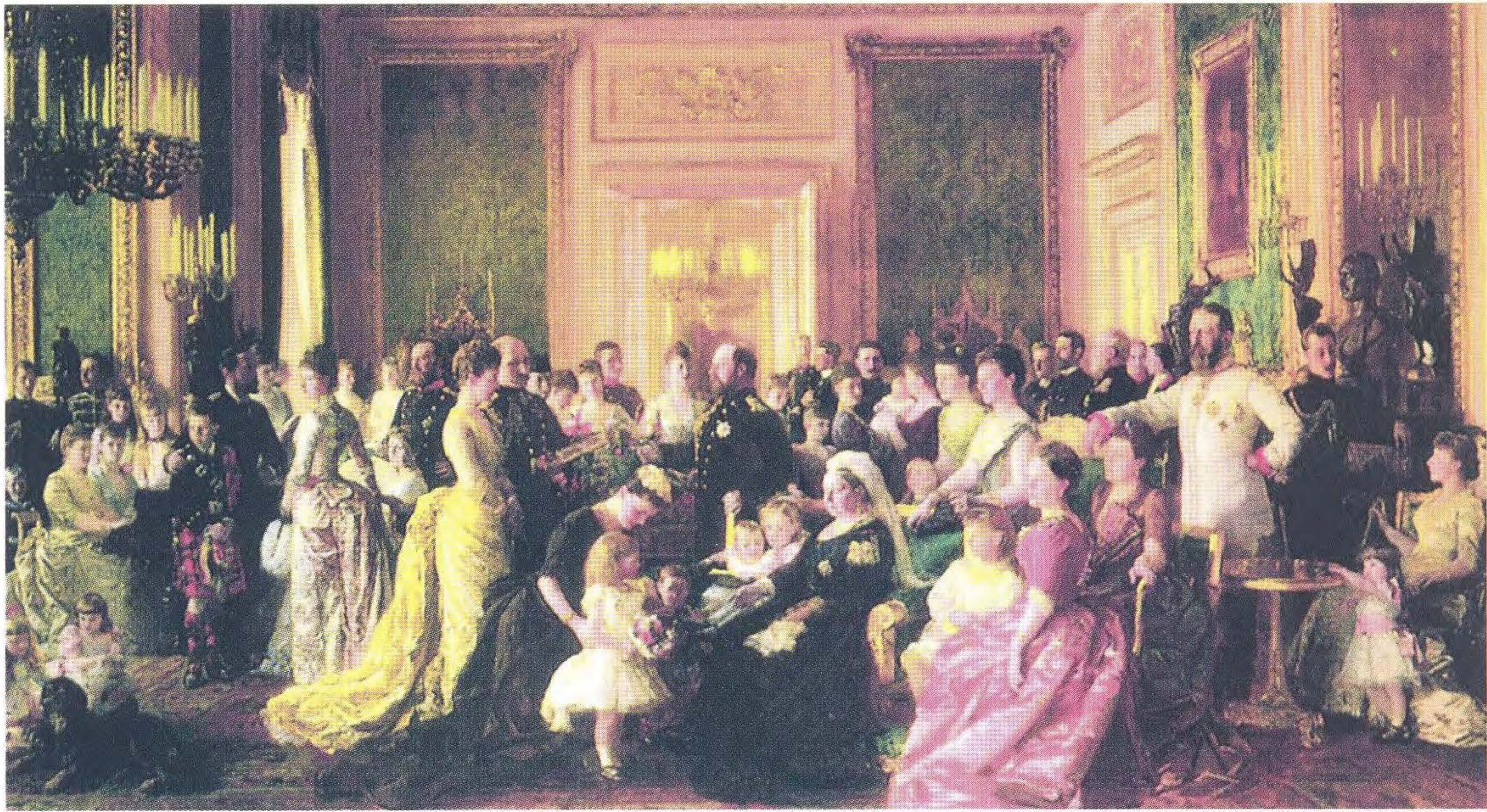
A distrofia muscular de Duchenne é uma doença recessiva fatal ligada ao X. O fenótipo é o definhamento e atrofia dos músculos. Geralmente, o início é antes da idade de 6 anos, com o confinamento a uma cadeira de rodas por volta dos 12 anos e morte aos 20. O gene da distrofia muscular de Duchenne codifica a proteína muscular distrofina. Esse conhecimento nos traz a esperança de uma melhor compreensão da fisiologia dessa condição e, finalmente, uma terapia.

Um raro fenótipo recessivo ligado ao X que é interessante do ponto de vista da diferenciação sexual é uma condição chamada de *síndrome de feminização testicular*, que tem uma frequência de cerca de 1 em 65.000 nascimentos masculinos. As pessoas afetadas por essa síndrome são cromossomicamente homens, tendo 44 autossomos mais um cromossomo X e um Y, mas desenvolvem-se como mulheres (Figura 2.38). Elas têm genitália externa masculina, uma vagina em fundo de saco e ausência de útero. Os testículos podem estar presentes seja nos grandes lábios, seja no abdome. Embora muitas dessas pessoas se casem, elas são estéreis. A condição não é revertida por tratamento com o hormônio masculino, sendo assim, às vezes, chamada de *síndrome de insensibilidade*

Herança de hemofilia na realeza européia



(a)



(b)

FIG. 2.37 Um heredograma para a condição recessiva ligada ao X da hemofilia nas famílias reais da Europa. Um alelo recessivo causador da hemofilia (falta de coagulação do sangue) surgiu nas células reprodutivas da rainha Victoria ou de um de seus genitores por mutação. Esse alelo se espalhou por outras famílias reais por meio de casamentos. (a) Esse heredograma parcial mostra homens afetados e mulheres portadoras (heterozigotas). A maioria dos que entraram para as famílias foram omitidos do heredograma por simplificação. Você pode deduzir a probabilidade de a atual família real ter o alelo recessivo? (b) Uma pintura mostrando a rainha Victoria e seus vários descendentes. [(a) C. Stern, *Principles of Human Genetics*, 3rd ed. Copyright 1973 by W.H. Freeman and Company. (b) "The Royal Collection © 2006, Her Majesty Queen Elizabeth II."]

androgênica. O motivo da insensibilidade é que a mutação no gene de receptor de andrógenos faz com que o receptor funcione mal, e, assim, o hormônio masculino não tem efeito nos órgãos-alvo que contribuem para a masculinidade. Em humanos, a feminilidade resulta quando o sistema determinante de masculinidade não é funcional.

Distúrbios dominantes ligados ao X

Os padrões de herança dos distúrbios dominantes ligados ao X têm as seguintes características nos heredogramas (Figura 2.39):

1. Os homens afetados passam a condição para todas as suas filhas, mas a nenhum de seus filhos.

Fenótipo de feminização testicular

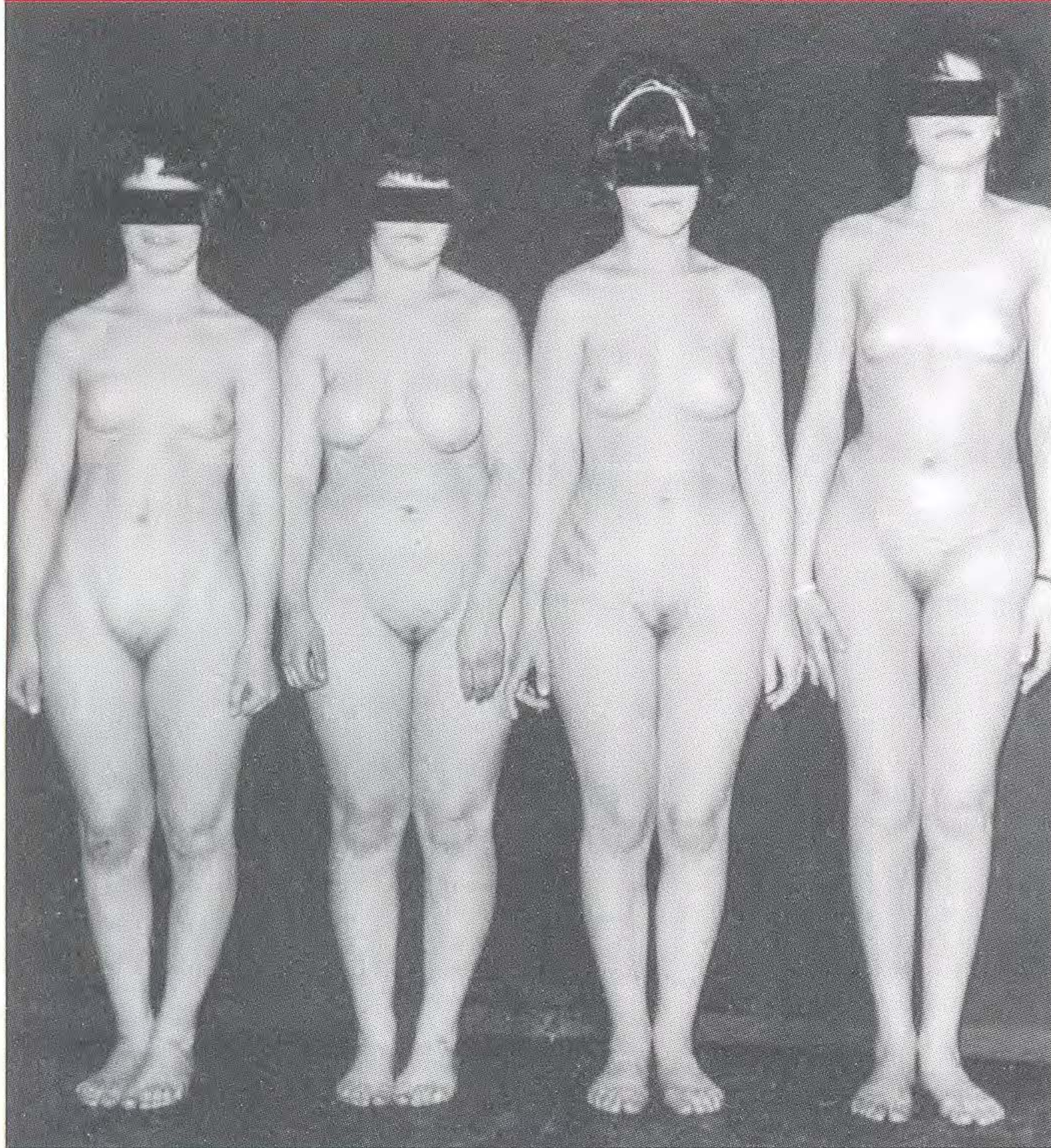


FIG. 2.38 As quatro irmãs nessa foto têm a síndrome de feminização testicular (insensibilidade congênita a andrógenos). Todas as quatro têm 44 autossomos mais um cromossomo X e um Y, mas herdaram um alelo recessivo ligado ao X que confere insensibilidade a andrógenos (hormônios masculinos). Uma de suas irmãs (não mostrada), que era geneticamente XX, era portadora e teve um filho que também tinha feminização testicular. [Leonard Pinsky, McGill University.]

2. As mulheres heterozigotas afetadas casadas com homens não-afetados passam a condição para metade de seus filhos e filhas.

Esse modo de herança não é comum. Um exemplo é a hipofosfatemia, um tipo de raquitismo resistente à vitamina D. Algumas formas de hipertricose (excesso de pêlos no corpo e face) apresentam herança dominante ligada ao X.

Herança ligada ao Y

Apenas os homens herdam genes na região diferencial do cromossomo humano Y, com os pais transmitindo os genes para seus filhos. O gene que tem um papel primário na masculinização é o **gene SRY**, às vezes chamado de *fator determinante testicular*. A análise genômica confirmou isso, e o gene SRY fica na região diferencial do cromossomo Y. Assim, a masculinização é ligada ao Y e apresenta o padrão esperado de transmissão exclusivamente de homem a homem. Alguns casos de esterilidade masculina foram demonstrados como sendo causados por deleções de regiões do cromossomo Y contendo genes que promovem espermatogênese. A esterilidade masculina não é herdável, mas, curiosamente, os pais desses homens têm cromossomos Y normais, mostrando que as deleções são novas.

Não têm ocorrido casos convincentes de variantes fenotípicas não-sexuais associadas ao cromossomo Y. Os pêlos nas orelhas (Figura 2.40) foram propostos como uma possibilidade, embora discutível. O fenótipo é extremamente raro entre as populações da maioria dos países, porém mais comum entre populações da Índia. Em algumas famílias, os pêlos nas orelhas foram demonstrados como sendo transmitidos exclusivamente de pais para filhos.

Os padrões de herança com representação desigual de fenótipos em homens e mulheres podem localizar os genes envolvidos com um dos cromossomos sexuais.

Herança de um distúrbio dominante ligado ao X

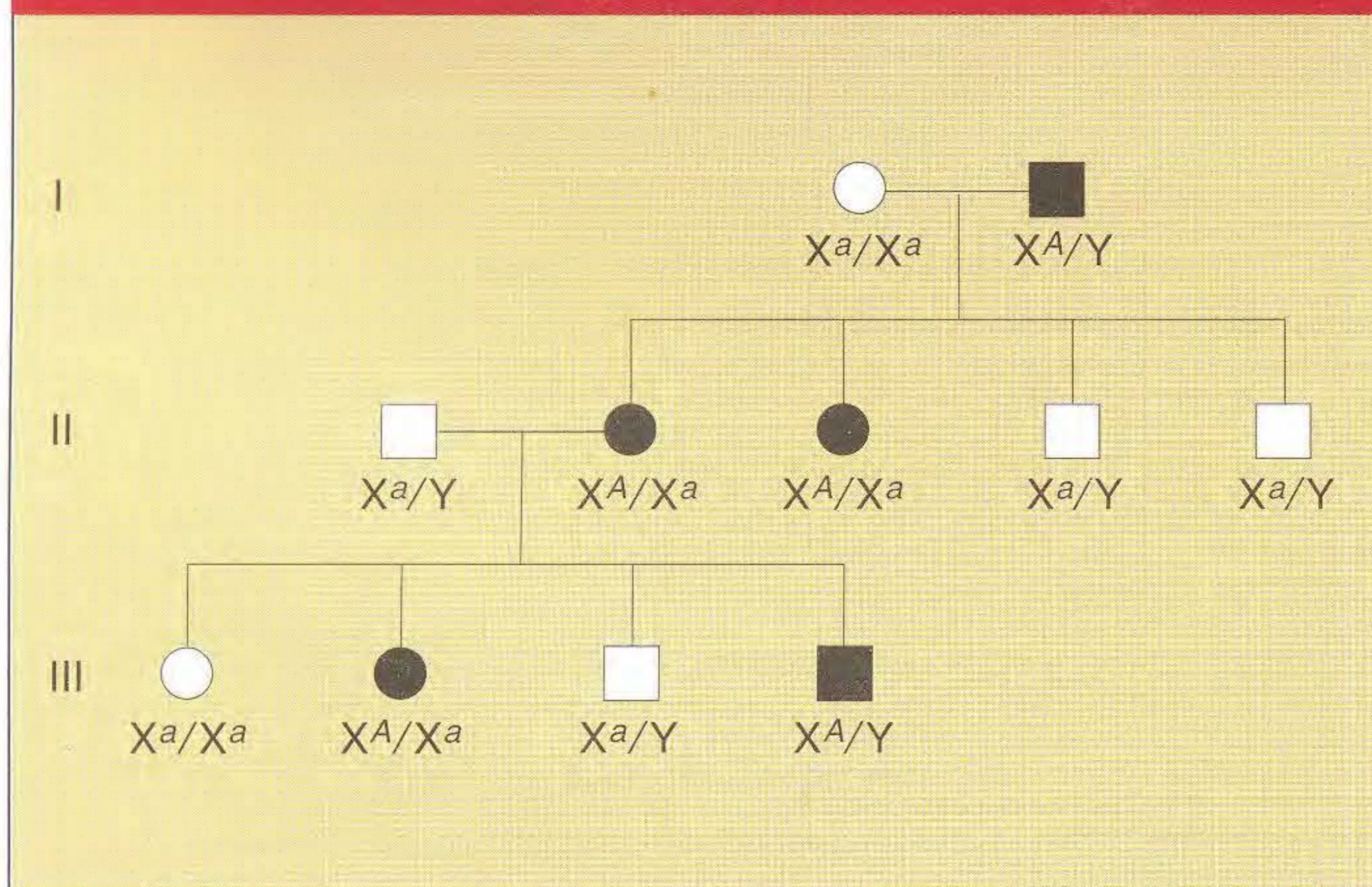


FIG. 2.39 Todas as filhas de um homem manifestando um fenótipo dominante ligado ao X apresentarão o fenótipo. As mulheres heterozigotas para um alelo dominante ligado ao X transmitirão a condição para metade de seus filhos e filhas.

Orelhas com pêlos: um fenótipo proposto como sendo ligado ao Y

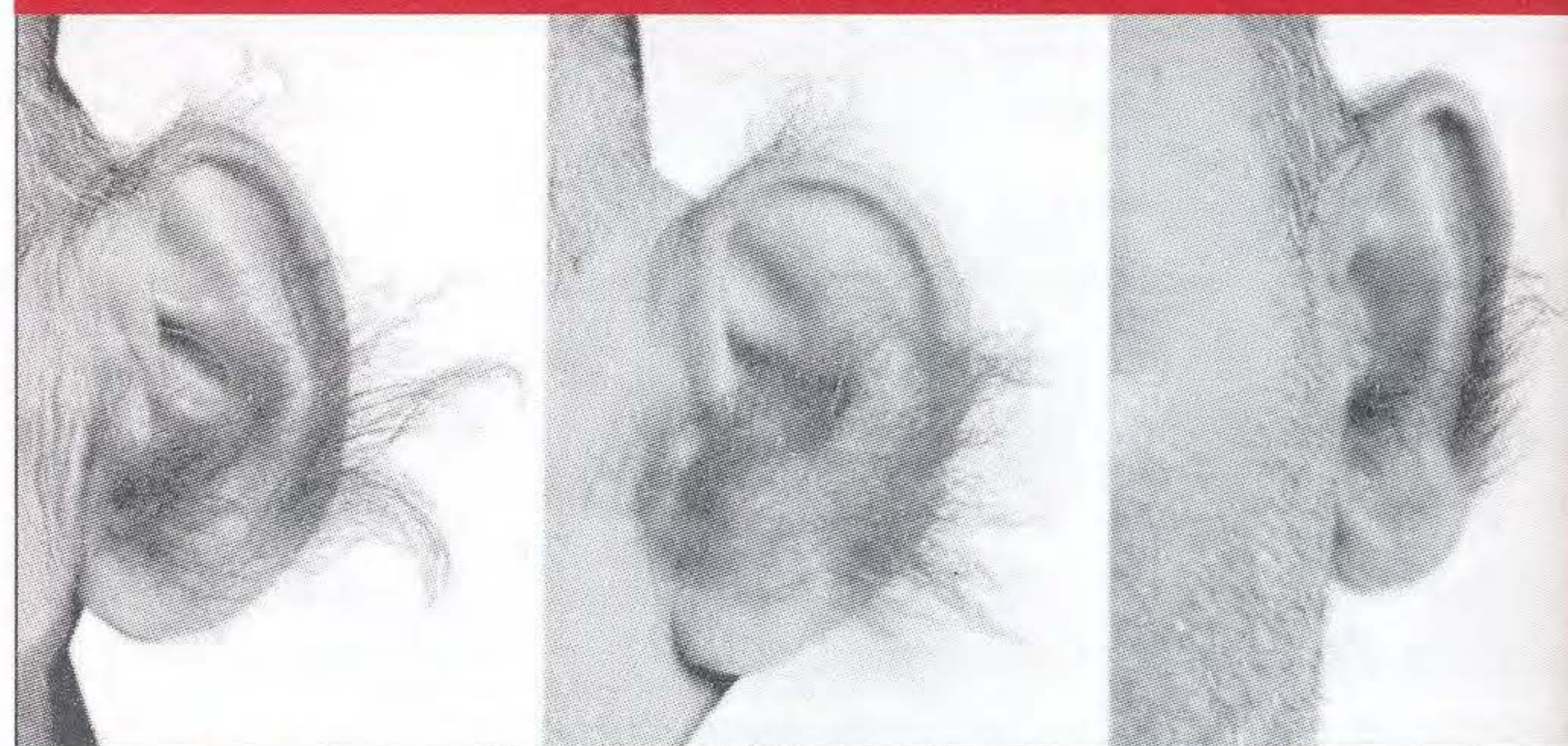
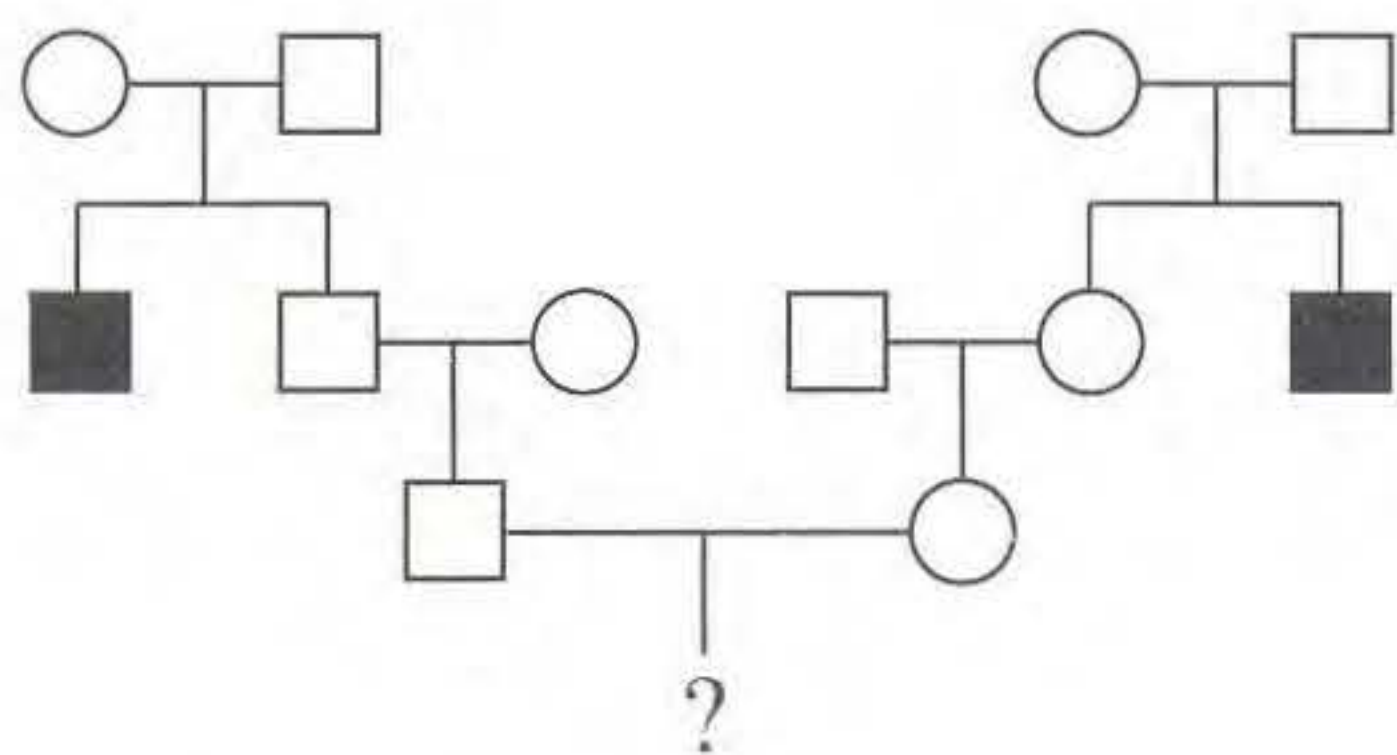


FIG. 2.40 Orelhas com pêlos foram propostas como sendo causadas por um alelo de um gene ligado ao Y. [De C. Stern, W.R. Centerwall, and S.S. Sarkar, *The American Journal of Human Genetics* 16, 1964, 467. Com permissão de Grune & Stratton, Inc.]

Cálculo de riscos na análise de heredogramas

Quando se sabe que um distúrbio com herança monogênica bem documentada está presente em uma família, o conhecimento dos padrões de transmissão podem ser usados para calcular a probabilidade de os prospectivos pais terem um filho com o distúrbio. Por exemplo, considere um caso no qual marido e esposa recém-casados descobrem que cada um tem um tio com doença de Tay-Sachs, uma grave doença autossômica recessiva causada por mau funcionamento da enzima hexosaminidase A. O defeito leva ao acúmulo de depósitos graxos nas células nervosas, causando paralisia seguida de morte precoce. O heredograma é o seguinte:



A probabilidade de o primeiro filho do casal ter Tay-Sachs pode ser calculada do seguinte modo. Como nenhum dos pais tem a doença, cada um só pode ser um homozigoto normal ou um heterozigoto. Se ambos forem heterozigotos, eles têm uma chance de passar o alelo recessivo para a criança, que então teria a doença de Tay-Sachs. Assim, devemos calcular a probabilidade de ambos serem heterozigotos, e, sendo assim, a probabilidade de passar o alelo deletério para a criança.

1. Os avós do marido devem ser ambos heterozigotos (T/t), pois eles produzem uma criança t/t (o tio). Portanto, eles efetivamente constituem uma reprodução monoíbrida. O pai do marido poderia ser T/T ou T/t , mas sabemos que as probabilidades relativas desses genótipos devem ser $1/4$ e $1/2$, respectivamente (a proporção esperada em uma reprodução monoíbrida é $\frac{1}{4} T/T$, $\frac{1}{2} T/t$ e $\frac{1}{4} t/t$). Portanto, há uma probabilidade de $2/3$ de que o pai seja um heterozigoto (dois terços é a proporção de prole não afetada de heterozigotos: $1/2$ dividida por $3/4$).

2. Supõe-se que a mãe do marido seja T/T , pois ela entrou para a família e os alelos da doença são geralmente raros. Assim, se o pai é T/t , então a reprodução com a mãe era $T/t \times T/T$ e as proporções esperadas na prole (que inclui o marido) são $\frac{1}{2} T/T$ e $\frac{1}{2} T/t$.
3. A probabilidade geral de o marido ser um heterozigoto deve ser calculada com o uso de uma regra estatística chamada **regra do produto**, que diz que

A probabilidade de dois eventos independentes ocorrerem é o produto de suas probabilidades individuais.

Como as transmissões gênicas em gerações diferentes são eventos independentes, podemos calcular a probabilidade de o marido ser um heterozigoto vezes a probabilidade de seu pai ter um filho heterozigoto, que é $2/3 \times 1/2 = 1/3$.

4. Do mesmo modo, a probabilidade de a esposa ser heterozigota também é $1/3$.
5. Se ambos forem heterozigotos (T/t), sua reprodução seria monoíbrida padrão, e, assim, a probabilidade de terem um filho t/t é $1/4$.
6. No geral, a probabilidade de o casal ter um filho afetado é a probabilidade de ambos serem heterozigotos e, então, ambos transmitirem o alelo recessivo para um filho. Novamente, esses eventos são independentes, e, assim, podemos calcular a probabilidade geral como $1/3 \times 1/3 \times 1/4 = 1/36$. Em outras palavras, há uma chance de 1 em 36 de eles terem um filho com Tay-Sachs.

Em algumas comunidades judaicas, o alelo Tay-Sachs não é tão raro quanto na população geral. Em tais casos, as pessoas não-afetadas que se casam em famílias com uma história de Tay-Sachs não podem ser supostas como sendo T/T . Se a frequência de heterozigotos T/t na comunidade é conhecida, essa frequência pode ser introduzida no cálculo da regra do produto. Hoje em dia, os testes de diagnóstico molecular para os alelos de Tay-Sachs estão disponíveis, e o uso cuidadoso desses testes reduziu acentuadamente a frequência da doença em algumas comunidades.

► Resumo

A análise do genoma de um organismo é um dos processos centrais da Genética. O genoma de um organismo eucariótico é composto de um conjunto de cromossomos de número e tamanho característicos desse organismo. Cada cromossomo contém uma molécula de DNA eficientemente helicoidizada por meio de proteínas histonas. Os genes são unidades transcritas ao longo dos cromossomos. O DNA situado entre os genes pode ser regulador ou de função desconhecida, sendo grande parte de elementos de transposição. As células diplóides contêm dois conjuntos de cromossomos idênticos (homólogos); as células haplóides contêm apenas um conjunto.

Na divisão das células somáticas, o genoma é transmitido por mitose, uma divisão nuclear. Nesse processo, cada cro-

mossomo replica-se em um par de cromátides e as cromátides são separadas para produzir duas células-filhas idênticas. (A mitose pode ocorrer em células diplóides ou haplóides.) Na meiose, que ocorre no ciclo sexual nos meiócitos, cada homólogo replica-se para formar uma díade de cromátides; então, as díades ficam pareadas formando uma tétrade, que segrega em cada uma das duas divisões celulares. O resultado é de quatro células haplóides, ou gametas. A meiose pode ocorrer apenas em uma célula diplóide, e organismos haplóides unem-se para formar um meiócito diplóide.

Um modo fácil de lembrar os principais eventos da meiose, usando seus dedos para representar os cromossomos, é mostrado na Figura 2.41.

Os principais eventos da mitose e meiose	
Mitose	Meiose
	
	
	
	
	

FIG. 2.41 Uso dos dedos para lembrar os principais eventos da mitose e da meiose.

A dissecção genética de uma propriedade biológica começa com uma coleção de mutantes. Cada mutante tem que ser testado para ver se ele é herdado como a mudança de um único gene. O procedimento seguido é essencialmente inalterado desde a época de Mendel, que fez a análise prototípica desse tipo. A análise é baseada em observar proporções fenotípicas específicas na prole de cruzamentos controlados. Em um caso típico, um cruzamento de $A/A \times a/a$ produz

uma F_1 que é toda A/a . Quando a F_1 é autofecundada ou entrecruzada, uma proporção genotípica de $\frac{1}{4} A/A : \frac{1}{2} A/a : \frac{1}{4} a/a$ é produzida na F_2 . (A nível fenotípico, esta proporção é $\frac{3}{4} A/- : \frac{1}{4} a/a$.) Os três genótipos monogênicos são homozigoto dominante, heterozigoto (monoíbrido) e homozigoto recessivo. Se um indivíduo A/a é cruzado com um a/a (um cruzamento-teste), é produzida uma proporção 1:1 na prole. As proporções 1:1, 3:1 e 1:2:1 vêm do princípio de segregação igual, pelo qual os produtos haplóides da meiose de A/a serão $\frac{1}{2} A$ e $\frac{1}{2} a$. A base celular da segregação igual de alelos é a segregação de cromossomos homólogos na meiose. Fungos haplóides podem ser usados para mostrar segregação igual ao nível de uma única meiose (uma proporção 1:1 em um asco).

A base molecular da produção de cromátide na meiose é a replicação do DNA. A segregação na meiose pode ser observada diretamente, a nível molecular (DNA), se forem usadas sondas adequadas para detectar os morfos de um dimorfismo molecular. A força molecular da segregação é a despolimerização e subsequente encurtamento de microtúbulos que são ligados aos centrômeros. As mutações recessivas são geralmente nos genes que são haplossuficientes, enquanto as mutações dominantes em geral são devidas à haploinsuficiência gênica.

Em muitos organismos, o sexo é determinado cromossomicamente, e, tipicamente, XX é fêmea e XY é macho. Os genes no cromossomo X (genes ligados ao X) não têm contraparte no cromossomo Y e apresentam um padrão de herança monogênica que difere nos dois sexos, em geral resultando em proporções diferentes de prole masculina e feminina.

A segregação monogênica mendeliana é útil na identificação de alelos mutantes subjacentes a muitos distúrbios humanos. A análise de heredogramas pode revelar distúrbios autossômicos ou ligados ao X dos tipos dominante e recessivo. A lógica da genética mendeliana tem que ser usada com cautela, levando em conta que o tamanho da prole humana é pequeno e as proporções fenotípicas não são necessariamente típicas das esperadas de tamanhos maiores de amostras. Se um distúrbio monogênico conhecido está presente em um heredograma, a lógica mendeliana pode ser usada para prever a possibilidade de os filhos herdarem a doença.

► Termos-chave

alelo	cromatina	diplóide
alelo nulo	cromossomo	dissecção genética
análise de heredogramas	cromossomo homólogo	dominante
asco	cromossomo sexual	elemento de transposição
autofecundação	cromossomo X	euromatina
banda cromossômica	cromossomo Y	fenótipo
bivalente	cruzamento	gene
característica	cruzamento monoíbrido	gene SRY
caráter	cruzamento-teste	genética direta
centrômero	descoberta do gene	genética reversa
cinetócoro	díade	genoma
cromátide	dimorfismo	genótipo

geração parental (P)	ligação ao Y	primeira geração filial (F ₁)
haplóide	linhagem pura	produto da meiose
haploinsuficiente	meiócito	propósito
haplossuficiente	meiose	propriedade
hemizigoto	mitose	recessivo
heterocromatina	monoíbrido	região pseudo-autossômica
heterozigoto	morfo	regra do produto
homólogo	mutação vazante	segunda geração filial (F ₂)
homozigoto	mutante	sexo heterogamético
homozigoto dominante	nucléolo	sexo homogamético
homozigoto recessivo	nucleossomo	telômero
íntron	número haplóide	testador
lei da segregação igual (Primeira Lei de Mendel)	organizador nucleolar	tétrade
ligação ao sexo	par de genes	tipo selvagem
ligação ao X	planta dióica	zigoto
	polimorfismo	

► Problemas Resolvidos

Esta seção em cada capítulo contém alguns problemas resolvidos que mostram como abordar os conjuntos de problemas seguintes. O propósito dos conjuntos de problemas é desafiar sua compreensão dos princípios genéticos aprendidos no capítulo. O melhor modo de demonstrar a compreensão de um assunto é ser capaz de usar esse conhecimento em uma situação real ou simulada. Esteja ciente de que não há um modo similar a uma máquina para resolver esses problemas. Os três recursos principais à sua disposição são os princípios genéticos aprendidos, lógica e ensaio-e-erro.

Aqui vai um aviso geral antes de começar. Primeiro, é absolutamente essencial ler e compreender todo o problema. A maioria dos problemas usa dados tirados de pesquisas que alguém fez: questione por que a pesquisa foi feita e qual era a provável meta. Descubra exatamente que fatos foram fornecidos, que suposições foram feitas, que indícios são dados no problema e que deduções podem ser feitas a partir da informação disponível. Segundo, seja metódico. Ficar olhando para o problema raramente ajuda. Refaça a informação no problema de seu próprio modo, de preferência usando uma representação diagramática ou um esquema para ajudá-lo a pensar sobre o problema. Boa sorte!

Problema resolvido 1. Foram feitos cruzamentos entre duas linhagens puras de coelhos que chamaremos de A e B. Um macho da linhagem A foi cruzado com uma fêmea da linhagem B, e os coelhos das F₁ foram subsequentemente intercruzados para produzir uma F₂. Três quartos dos animais da F₂ apresentaram uma gordura subcutânea branca, e um quarto, gordura subcutânea amarela. Depois, a F₁ foi examinada e apresentava gordura branca. Vários anos depois, foi feita uma tentativa de repetir o experimento usando o mesmo macho da linhagem A e a mesma fêmea da linhagem B. Dessa vez, a F₁ e toda a F₂ (22 animais) tinham gordura branca. A única diferença entre o experimento original e o repetido que parecia relevante foi que, no original, todos os animais eram alimentados com vegetais frescos, enquanto, na repetição, eram alimentados com ração comercial. Forneça uma explicação para a diferença e um teste para sua idéia.

SOLUÇÃO

Da primeira vez que o experimento foi feito, os criadores estavam perfeitamente certos em propor que um par de alelos determina a gordura corpórea branca *versus* amarela porque os dados claramente lembravam os resultados de Mendel em ervilhas. Branco deve ser dominante, e, assim, podemos representar o alelo branco como *W* e o alelo amarelo como *w*. Os resultados podem, então, ser expressos do seguinte modo:

P	$W/W \times w/w$
F ₁	w/w
F ₂	$\frac{1}{4} W/W$
	$\frac{1}{2} W/w$
	$\frac{1}{4} w/w$

Sem dúvida, se os coelhos parentais tivessem sido sacrificados, um genitor (não podemos dizer qual) seria previsto tendo gordura branca e o outro amarela. Felizmente, os coelhos não foram sacrificados, e os mesmos animais foram novamente cruzados, levando a um resultado diferente muito interessante. Em geral, em ciência, uma observação inesperada pode levar a um novo princípio, e, em vez de passar para outra coisa, é útil tentar explicar a inconsistência. Por que a proporção 3:1 desapareceu? Aqui vão algumas explicações possíveis.

Primeiro, talvez os genótipos dos animais parentais tenham mudado. Esse tipo de mudança espontânea afetando todo o animal, ou pelo menos suas gônadas, é muito improvável, porque mesmo a experiência comum nos diz que os organismos tendem a ficar estáveis em seu tipo.

Segundo, na repetição, a amostra de 22 animais F₂ não contém uma gordura amarela apenas por acaso (“azar”). Essa explicação, novamente, parece improvável, pois a amostra é bem grande, mas é uma possibilidade.

Uma terceira explicação refere-se ao princípio coberto no Capítulo 1 de que os genes não atuam no vácuo; eles dependem do ambiente para seus efeitos. Assim, a fórmula “genótipo + ambiente = fenótipo” é mnemonicamente útil.

Um corolário dessa fórmula é que os genes podem atuar diferentemente em ambientes diferentes; assim,

$$\text{genótipo 1} + \text{ambiente 1} = \text{fenótipo 1}$$

e

$$\text{genótipo 1} + \text{ambiente 2} = \text{fenótipo 2}$$

No problema em questão, as dietas diferentes constituíram ambientes diferentes, e, assim, uma possível explicação dos resultados é que o alelo recessivo w produz gordura amarela só quando a dieta contém vegetais frescos. Essa explicação é testável. Um modo de testá-la é repetir o experimento e usar vegetais como alimento, mas os genitores podem estar mortos nessa época. Um modo mais convincente é cruzar vários dos coelhos F_2 de gordura branca do segundo experimento. De acordo com a interpretação original, alguns deles devem ser heterozigotos, e, se sua prole for criada com vegetais, deve aparecer a gordura amarela em proporções mendelianas. Por exemplo, se um cruzamento é W/w e w/w , a prole seria $\frac{1}{2}$ de gordura branca e $\frac{1}{2}$ de gordura amarela.

Se não ocorrer esse resultado e não aparecer prole tendo gordura amarela em nenhum dos cruzamentos, seríamos forçados de volta à primeira ou segunda explicação. A segunda explicação pode ser testada usando números maiores, e, se essa explicação não funcionar, ficaremos com a primeira explicação, que é difícil de testar diretamente.

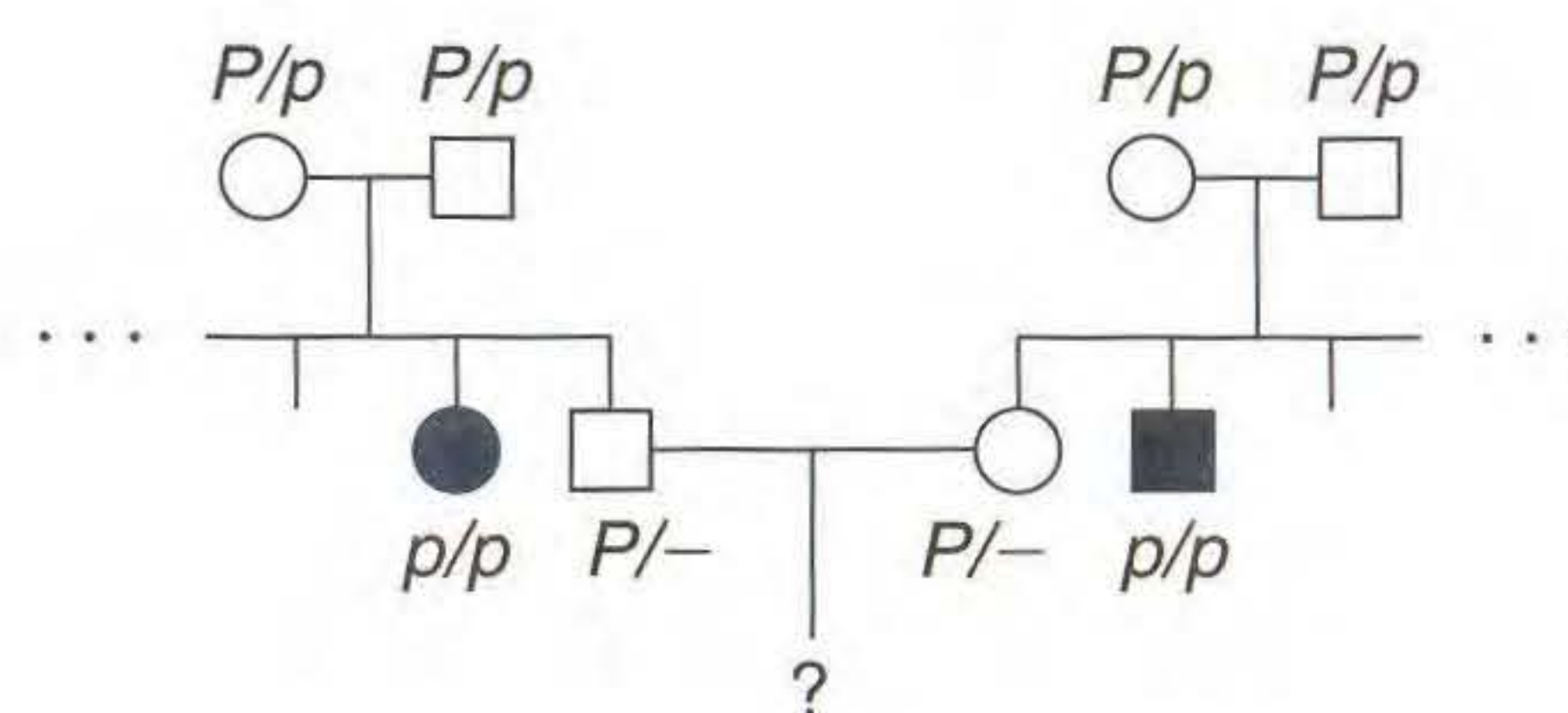
Como você deve ter suspeitado, na realidade a dieta é que é a responsável. Os detalhes específicos ilustram os efeitos ambientais muito bem. Os vegetais frescos contêm substâncias amarelas chamadas xantofilas, e o alelo dominante W dá aos coelhos a habilidade em degradar essas substâncias para uma forma incolor ("branca"). Entretanto, os animais w/w não têm essa habilidade, e as xantofilas são depositadas na gordura, tornando-a amarela. Quando não foram ingeridas xantofilas, ambos os animais $W/-$ e w/w acabam tendo gordura branca.

Problema resolvido 2. A fenilcetonúria é uma doença hereditária humana que resulta da incapacidade do corpo em processar a substância fenilalanina, que é contida nas proteínas que comemos. A PKU manifesta-se na lactância e, se não tratada, geralmente leva ao retardo mental. A PKU é causada por um alelo recessivo com herança mendeliana simples.

Um casal pretende ter um filho, mas procura um consultor genético porque o homem tem uma irmã com PKU e a mulher tem um irmão com PKU. Não existem outros casos conhecidos em suas famílias. Eles pedem ao consultor genético que determine a probabilidade de que seu primeiro filho tenha PKU. Qual é essa probabilidade?

SOLUÇÃO

O que podemos deduzir? Chamemos o alelo causador do fenótipo PKU de p e o respectivo alelo normal de P ; então, a irmã e o irmão do homem e da mulher, respectivamente, devem ser p/p . Para produzir essas pessoas afetadas, todos os quatro avós devem ter sido heterozigotos normais. O heredograma pode ser resumido assim:



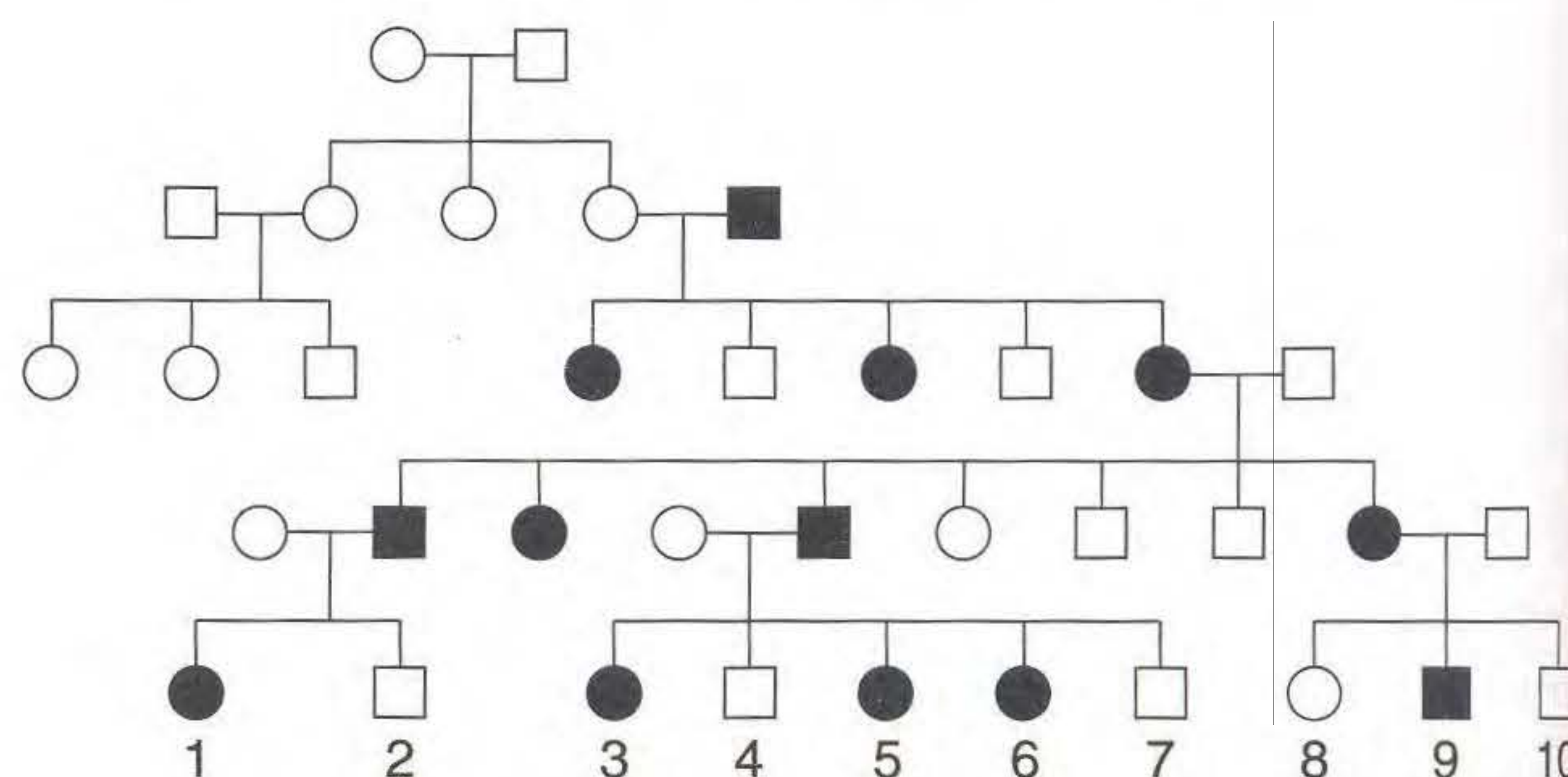
Uma vez feitas essas deduções, o problema é reduzido a uma aplicação da regra do produto. O único modo pelo qual o homem e a mulher podem ter um filho PKU é se ambos forem heterozigotos (é óbvio que eles mesmos não têm a doença). Ambas as reproduções dos avós são monoíbridas mendelianas simples, esperando-se que produzam uma prole nas seguintes proporções:

$$\begin{array}{l} \frac{1}{4} P/P \\ \frac{1}{2} P/p \\ \frac{1}{4} p/p \end{array} \quad \begin{array}{l} \text{Normal } (\frac{3}{4}) \\ \text{PKU } (\frac{1}{4}) \end{array}$$

Sabemos que o homem e a mulher são normais, e, assim, a probabilidade de cada um ser heterozigoto é $\frac{2}{3}$ porque, dentro da classe $P/-$, $\frac{2}{3}$ são P/p e $\frac{1}{3}$ P/P .

A probabilidade de *tanto* o homem quanto a mulher serem heterozigotos é $\frac{2}{3} \times \frac{2}{3} = \frac{4}{9}$. Se ambos são heterozigotos, então um quarto de seus filhos teria PKU, e, assim, a probabilidade de que seu primeiro filho tenha PKU é $\frac{1}{4}$, e a probabilidade de serem heterozigotos e seu primeiro filho ter PKU é $\frac{4}{9} \times \frac{1}{4} = \frac{4}{36} = \frac{1}{9}$, que é a resposta.

Problema resolvido 3. Uma doença humana rara afetou uma família como mostrado no heredograma seguinte.



a. Deduza o modo mais provável de herança.

b. Quais seriam os resultados dos casamentos de primos 1×9 , 1×4 , 2×3 e 2×8 ?

SOLUÇÃO

a. O modo mais provável de herança é dominante ligada ao X. Supomos que o fenótipo da doença é dominante porque, após ser introduzido no heredograma pelo homem da geração II, ele aparece em todas as gerações. Supomos que o fenótipo é ligado ao X porque os pais não o transmitem para seus filhos. Se fosse autossômico dominante, a transmissão de pai para filho seria comum.

Na teoria, a autossômica recessiva poderia funcionar, mas é improvável. Em particular, note os casamentos entre os membros afetados da família e os não-afetados de fora dela. Se a condição fosse autossômica recessiva, o único modo pelo qual esses casamentos poderiam ter prole afetada seria se cada pessoa que entra para a família fosse um heterozigoto; então, os casamentos seriam a/a (afetado) \times A/a (não-afetado). Entretanto, sabemos que a doença é rara. Em tal caso, é altamente improvável os heterozigotos serem comuns. A herança recessiva ligada ao X é impossível, pois uma reprodução de uma mulher afetada com um homem normal não produziria filhas afetadas. Faremos A representar o alelo causador da doença e a representar o alelo normal.

b. 1×9 : Número 1 deve ser heterozigoto A/a porque ela deve ter obtido a de sua mãe normal. O número 9 deve ser A/Y . Assim, a reprodução é $\text{♀ } A/a \times \text{♂ } A/Y$.

Gametas femininos	Gametas masculinos	Prole
$\frac{1}{2} A$	$\frac{1}{2} A$	$\frac{1}{4} A/A \text{ ♀}$
	$\frac{1}{2} Y$	$\frac{1}{4} A/Y \text{ ♂}$
$\frac{1}{2} a$	$\frac{1}{2} A$	$\frac{1}{4} A/a \text{ ♀}$
	$\frac{1}{2} Y$	$\frac{1}{4} a/Y \text{ ♂}$

1×4 : Deve ser $\text{♀ } A/a \times \text{♂ } a/Y$.

Gametas femininos	Gametas masculinos	Prole
$\frac{1}{2} A$	$\frac{1}{2} a$	$\frac{1}{4} A/a \text{ ♀}$
	$\frac{1}{2} Y$	$\frac{1}{4} A/Y \text{ ♂}$
$\frac{1}{2} a$	$\frac{1}{2} a$	$\frac{1}{4} a/a \text{ ♀}$
	$\frac{1}{2} Y$	$\frac{1}{4} a/Y \text{ ♂}$

2×3 : Deve ser $\text{♂ } a/Y \times \text{♀ } A/a$ (mesmo que 1×4).

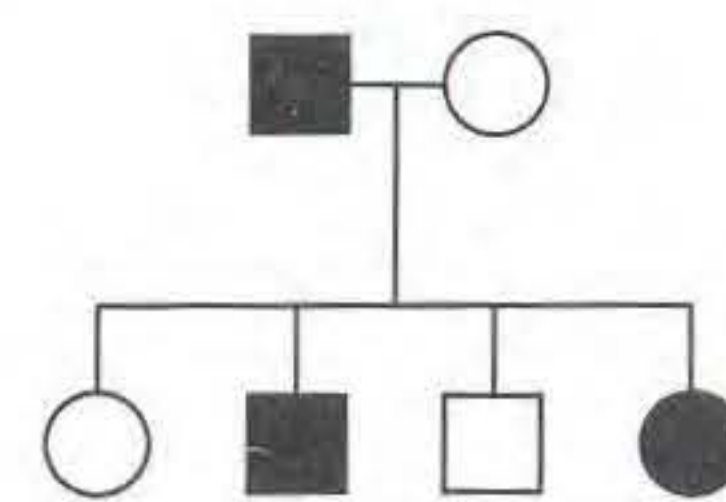
2×8 : Deve ser $\text{♂ } a/Y \times \text{♀ } a/a$ (toda prole normal).

► Problemas

PROBLEMAS BÁSICOS

1. Faça uma frase incluindo as palavras *cromossomo*, *genes* e *genoma*.
2. As ervilhas são diplóides e $2n = 14$. *Neurospora* é um fungo haplóide e $n = 7$. Se fosse possível fracionar o DNA genômico de ambos usando eletroforese de campo pulsado, quantas bandas distintas de DNA seriam visíveis em cada espécie?
3. O feijão (*Vicia faba*) é diplóide e $2n = 18$. Cada conjunto cromossômico haplóide contém aproximadamente 4 m de DNA. O tamanho médio de cada cromossomo, durante a metáfase da mitose, é de 13 μm . Qual a proporção média de compactação do DNA na metáfase? (Proporção de compactação = tamanho do cromossomo/tamanho da molécula de DNA.) Como é obtida essa compactação?
4. Se chamarmos a quantidade de DNA por genoma de " x ", cite uma situação ou situações em organismos diplóides nas quais a quantidade de DNA por célula é
 - a. x
 - b. $2x$
 - c. $4x$
5. Cite a principal função da mitose.
6. Cite as duas principais funções da meiose.
7. Você pode criar um sistema diferente de divisão nuclear que atinja o mesmo resultado que o da meiose?
8. Em um possível cenário futuro, a fertilidade masculina vai a zero, mas, felizmente, os cientistas desenvolvem um modo para as mulheres produzirem crianças por nascimento virgem. Os meiócitos são convertidos diretamente (sem sofrer meiose) em zigotos, que se implantam do modo usual. Quais seriam os efeitos em curto e longo prazos em tal sociedade?
9. De que modos a segunda divisão da meiose difere da mitose?
10. Faça um recurso mnemônico para lembrar os cinco estágios da prófase I da meiose e os quatro estágios da mitose.
11. Em uma tentativa de simplificar a meiose para o benefício dos estudantes, cientistas loucos desenvolveram um modo de evitar a fase S pré-meiótica e fazer com que ocorra só uma divisão, incluindo pareamento, crossing over e segregação. Esse sistema funcionaria, e os produtos de tal sistema seriam diferentes dos do sistema atual?
12. Theodor Boveri disse: "O núcleo não se divide; ele é dividido." O que ele queria dizer?
13. Francis Galton, um geneticista da era pré-mendeliana, percebeu o princípio de que metade de nossa constituição genética é derivada de cada genitor, um quarto de cada avô e um oitavo de cada bisavô, e assim por diante. Ele estava certo? Explique.
14. Se as crianças recebem metade de seus genes de um genitor e metade do outro genitor, porque os irmãos não são idênticos?
15. Diga onde as células se dividem mitoticamente e onde se dividem meioticamente em uma samambaia, um musgo, uma planta com flores, um pinheiro, um cogumelo, um sapo, uma borboleta e um caramujo.
16. As células humanas normalmente têm 46 cromossomos. Para cada um dos seguintes estágios, diga o número de moléculas de DNA nuclear presentes em uma célula humana:

- a. Metáfase da mitose
b. Metáfase I da meiose
c. Telófase da mitose
d. Telófase I da meiose
e. Telófase II da meiose
17. Quatro dos seguintes eventos são parte tanto da meiose quanto da mitose, mas apenas um é meiótico. Qual deles? (1) Formação de cromátides, (2) formação do fuso, (3) condensação cromossômica, (4) movimento cromossômico para os pólos, (5) sinapse.
18. Em milho, o alelo f' causa um endosperma farinhento e um alelo f'' causa um endosperma pétreo. No cruzamento $\text{♀ } f'/f' \times \text{♂ } f''/f''$, todos os endospermas da prole são farinhentos, mas, no cruzamento recíproco, todos os endospermas da prole são pétreos. Qual a explicação possível? (Veja legenda da Figura 2.14.)
19. Qual é a Primeira Lei de Mendel?
20. Se você tivesse uma mosca-das-frutas (*Drosophila melanogaster*) que fosse de fenótipo A, que teste você faria para determinar se o genótipo da mosca era A/A ou A/a ?
21. Examinando uma grande amostra de colônias de leveduras em uma placa de Petri, um geneticista descobre uma colônia de formato anormal que é muito pequena. Essa colônia pequena foi cruzada com o tipo selvagem, e os produtos da meiose (ascosporos) foram espalhados em uma placa para produzir colônias. No total, havia 188 colônias tipo selvagem (tamanho normal) e 180 pequenas.
- a. O que pode ser deduzido desses resultados quanto à herança do fenótipo colônia pequena? (Invente símbolos genéticos.)
b. Com que se pareceria um asco desse cruzamento?
22. Dois porquinhos-da-índia (cobaias) foram cruzados e, durante vários anos, produziram uma prole com 29 pretos e 9 brancos. Explique esses resultados, dando os genótipos dos genitores e prole.
23. Em um fungo com quatro ascosporos, um alelo mutante $lis-5$ faz com que os ascosporos que possuam o alelo sejam brancos, enquanto o alelo tipo selvagem $lis-5^+$ resulta em ascosporos pretos. (Ascospores são esporos que constituem os quatro produtos da meiose.) Desenhe um asco de cada um dos seguintes cruzamentos:
a. $lis-5 \times lis-5^+$
b. $lis-5 \times lis-5$
c. $lis-5^+ \times lis-5^+$
24. Para determinado gene em um organismo diplóide, são necessárias oito unidades de produtos proteicos para o funcionamento normal. Cada alelo tipo selvagem produz cinco unidades.
- a. Se uma mutação cria um alelo nulo, você acha que esse alelo será recessivo ou dominante?
b. Que suposições devem ser feitas para responder à parte a?
25. Uma colônia de *Neurospora* na margem de uma placa parecia ser esparsa (baixa densidade) em comparação a outras colônias na placa. Essa colônia foi considerada um possível mutante, e, assim, foi removida e cruzada com um tipo selvagem do tipo reprodutivo oposto. Desse cruzamento foi obtida uma prole de 100 ascosporos. Nenhuma das colônias desses ascosporos era esparsa; todas pareciam ser normais. Qual a explicação mais simples desse resultado? Como você testaria sua explicação? (Nota: *Neurospora* é haplóide.)
26. De uma triagem em grande escala de muitas plantas de *Collinsia grandiflora*, foi descoberta uma planta com três cotilédones (normalmente, existem dois cotilédones). Essa planta foi cruzada com uma planta tipo selvagem normal pura, e 600 sementes desse cruzamento foram plantadas. Havia 298 plantas com dois cotilédones e 302 com três cotilédones. O que pode ser deduzido sobre a herança de três cotilédones? Invente símbolos genéticos como parte de sua explicação.
27. Na planta *Arabidopsis thaliana*, um geneticista está interessado no desenvolvimento de tricomas (pequenas projeções). Uma grande triagem descobriu duas plantas mutantes (A e B) que não tinham tricomas, e esses mutantes pareciam ser potencialmente úteis no estudo do desenvolvimento de tricomas. (Se eles foram determinados por mutações monogênicas, então descobrir as funções normais e anormais desses genes seria instrutivo.) Cada planta é cruzada com o tipo selvagem; em ambos os casos, as gerações seguintes (F_1) tinham tricomas normais. Quando as plantas de F_1 foram autofecundadas, a F_2 resultante foi a seguinte:
- F_2 do mutante A: 602 normais; 198 não-tricomas
 F_2 do mutante B: 267 normais; 93 não-tricomas
- a. O que esses resultados mostram? Inclua os genótipos propostos de todas as plantas em sua resposta.
b. Com base em sua explicação à parte a, é possível prever com confiança a F_1 do cruzamento do mutante original A com o mutante original B?
28. Você tem três dados: um vermelho (R), um verde (G) e um azul (B). Quando todos os três dados são lançados ao mesmo tempo, calcule a probabilidade dos seguintes resultados:
a. 6 (R), 6 (G), 6 (B)
b. 6 (R), 5 (G), 6 (B)
c. 6 (R), 5 (G), 4 (B)
d. Nenhum seis
e. Um número diferente em todos os dados
29. No heredograma a seguir, os símbolos pretos representam indivíduos com uma doença sanguínea muito rara.



Se você não tem nenhuma outra informação para continuar, você acha que a doença mais provavelmente é dominante ou recessiva? Cite seus raciocínios.

30. a. A habilidade em sentir o gosto da substância feniltiocarbamida é um fenótipo autossômico dominante, e a incapacidade é recessiva. Se uma mulher sensível se casa com um homem não-sensível que, em um casamento

anterior, teve uma filha não-sensível, qual a probabilidade de que seu primeiro filho seja

- (1) uma menina não-sensível
- (2) uma menina sensível
- (3) um menino sensível

b. Qual a probabilidade de que seus primeiros dois filhos sejam sensíveis de qualquer sexo?

- 31.** John e Martha estão pensando em ter filhos, mas o irmão de John tem galactosemia (uma doença autossômica recessiva) e a bisavó de Martha também teve galactosemia. Martha tem uma irmã que tem três filhos, nenhum dos quais tem galactosemia. Qual a probabilidade de que o primeiro filho de John e Martha tenha galactosemia?



Solucionando o Problema 31

1. O problema pode ser reformulado como um heredograma? Caso sim, faça um.
2. Partes do problema podem ser refeitas usando os quadrados de Punnett?
3. Partes do problema podem ser refeitas usando diagramas ramificados?
4. No heredograma, identifique uma reprodução que ilustre a Primeira Lei de Mendel.
5. Defina todos os termos científicos no problema, e procure outros termos sobre os quais você não tem certeza.
6. Que suposições precisam ser feitas para responder a este problema?
7. Que membros não-mencionados da família devem ser considerados? Por quê?
8. Que regras estatísticas devem ser relevantes, e em que situações elas podem ser aplicadas? Tais situações existem no problema?
9. Quais as duas generalidades sobre doenças autossômicas recessivas em populações humanas?
10. Qual a relevância da raridade do fenótipo em estudo na análise do heredograma em geral, e o que pode ser deduzido neste problema?
11. Nessa família, quais genótipos são certos e quais são incertos?
12. De que modo o lado de John do heredograma é diferente do lado de Martha? Como essa diferença afeta seus cálculos?
13. Existe alguma informação irrelevante no problema?
14. De que modo resolver esse tipo de problema é similar a resolver problemas que você já resolveu bem-sucedidamente? De que modo ele é diferente?
15. Você pode escrever uma história curta com base no dilema humano deste problema?

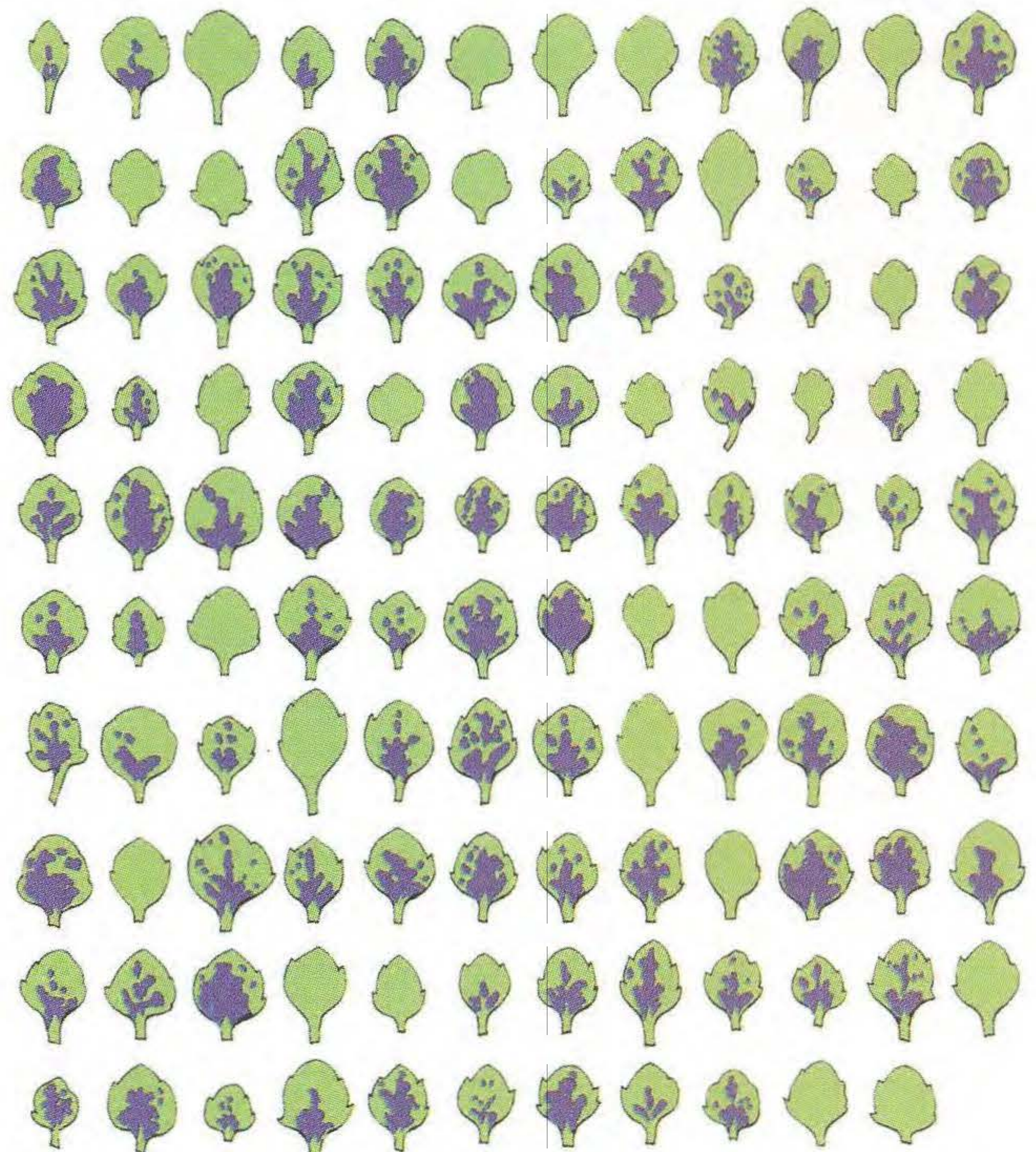
Agora tente resolver o problema. Se você não for capaz de fazê-lo, tente identificar o obstáculo e escreva uma frase ou duas descrevendo sua dificuldade. Então volte aos itens anteriores e veja se algum deles está relacionado à sua dificuldade.

- 32.** O gado Holstein normalmente é preto e branco. Um ótimo touro preto e branco, Charlie, foi comprado por um

fazendeiro por US\$100.000. Toda a prole de Charlie era de aspecto normal. Entretanto, alguns pares de sua prole, quando entrecruzados, produziram prole vermelha e branca com uma frequência de cerca de 25%. Charlie foi logo removido das listas de procriadores Holstein. Use símbolos para explicar exatamente o motivo.

- 33.** Suponha que o marido e a mulher são ambos heterozigotos para um alelo recessivo para albinismo. Se eles tiverem gêmeos dizigóticos (dois ovócitos), qual a probabilidade de que ambos os gêmeos tenham mesmo fenótipo de pigmentação?

- 34.** A planta canchalágua cresce na ilha de Vancouver e na parte continental da British Columbia. As populações são dimórficas para manchas púrpura nas folhas. Algumas plantas têm manchas e outras não. Perto de Nanaimo, uma planta na natureza tinha folhas manchadas. Essa planta, que ainda não tem flores, foi colhida e levada para o laboratório, onde foi deixada para autofecundar-se. As sementes foram coletadas e cresceram em uma prole. Uma folha colhida aleatoriamente (mas típica) de cada uma da prole é mostrada na ilustração acompanhante.

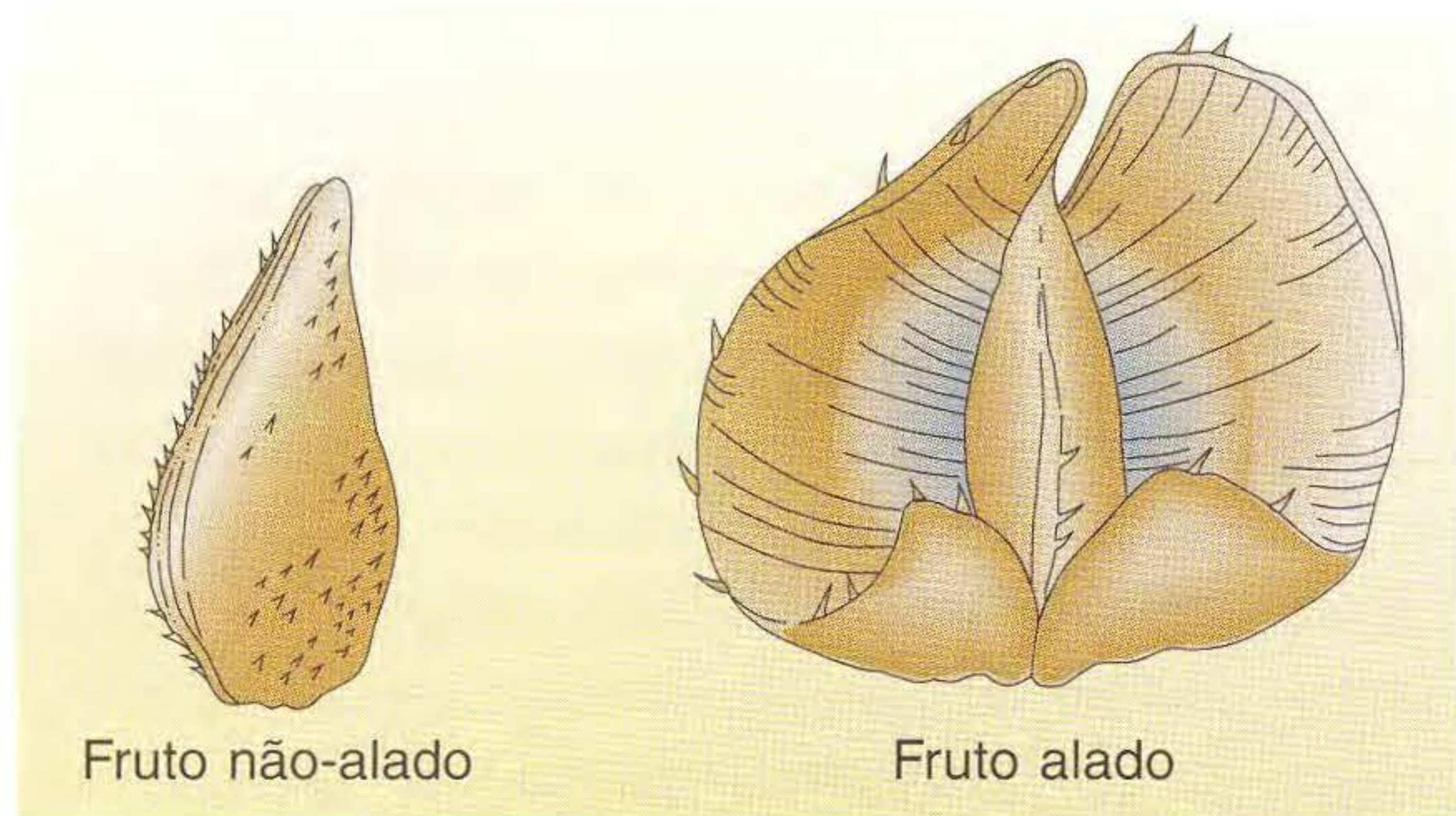


a. Formule uma hipótese genética concisa para explicar esses resultados. Explique todos os símbolos e mostre todas as classes genotípicas (e o genótipo da planta original).

b. Como você testaria sua hipótese? Seja específico.

- 35.** Pode ser provado que um animal não é portador de um alelo recessivo (isto é, não é heterozigoto para determinado gene)? Explique.

- 36.** Na natureza, a planta *Plectritis congesta* é dimórfica para a forma do fruto; isto é, as plantas têm frutos sem asas



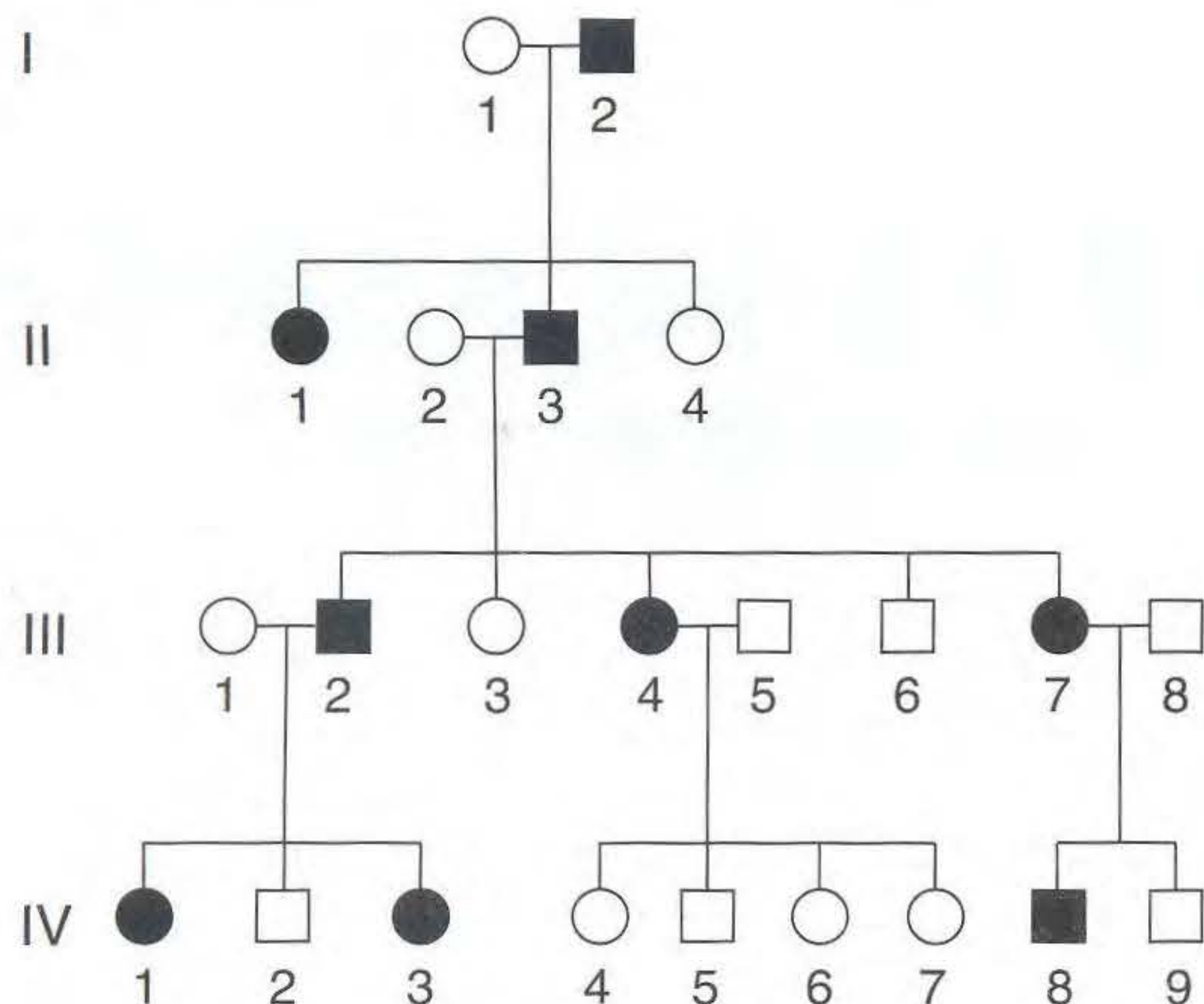
ou alados, como mostrado na ilustração. Foram coletadas plantas da natureza antes da floração e foram cruzadas ou autofecundadas com os seguintes resultados:

Polinização	Número de prole	
	Alados	Não-alados
Alados (autofecundados)	91	1*
Alados (autofecundados)	90	30
Não-alados (autofecundados)	4*	80
Alados × não-alados	161	0
Alados × não-alados	29	31
Alados × não-alados	46	0
Alados × alados	44	0
Alados × alados	24	0

*O fenótipo provavelmente tem uma explicação monogênica.

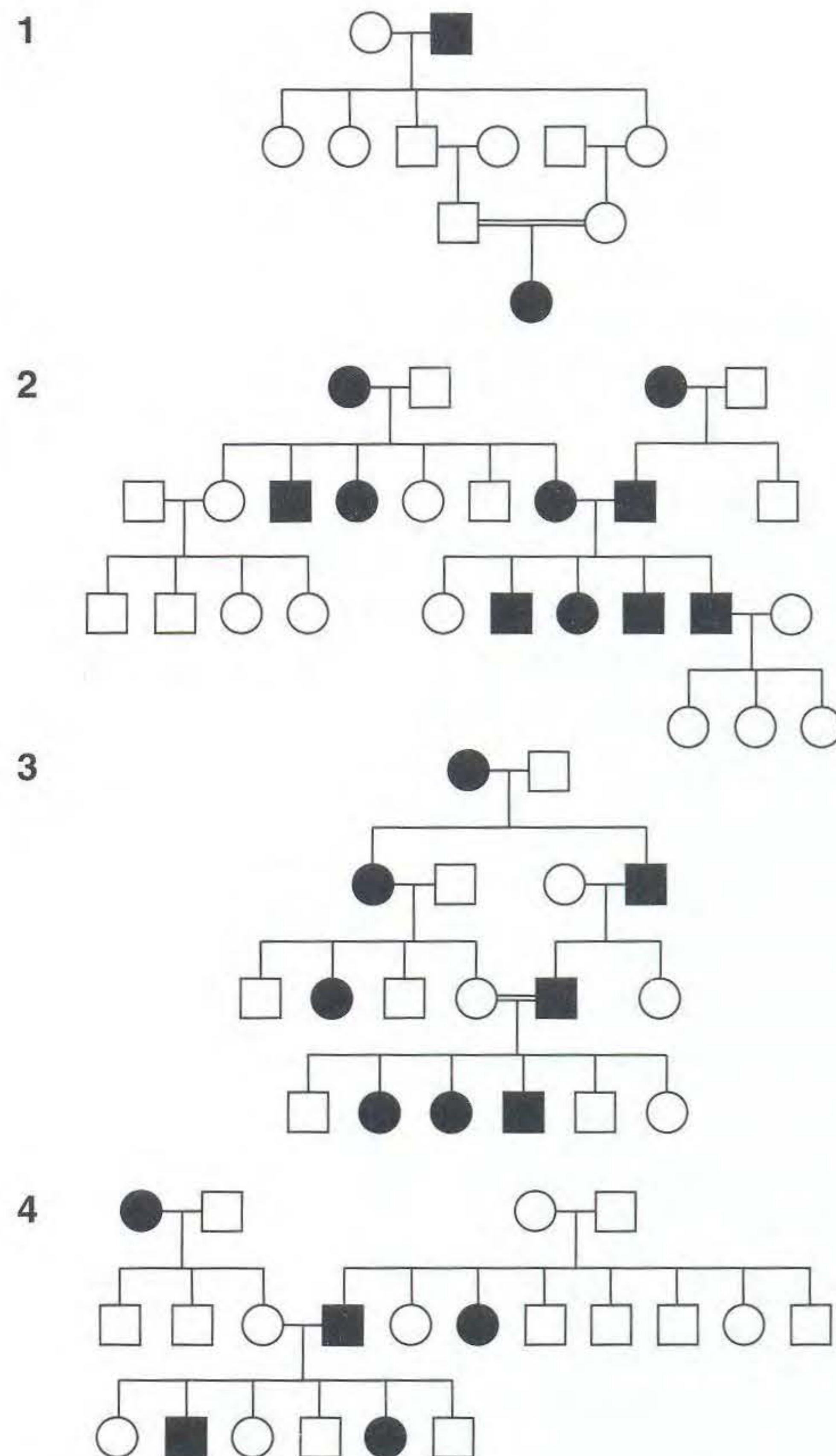
Interprete esses resultados e deduza o modo de herança desses fenótipos em forma de fruto. Use símbolos. Qual você imagina que seja a explicação não-genética para os fenótipos marcados com asteriscos no quadro?

37. O heredograma seguinte é de um raro mas relativamente brando distúrbio hereditário de pele.

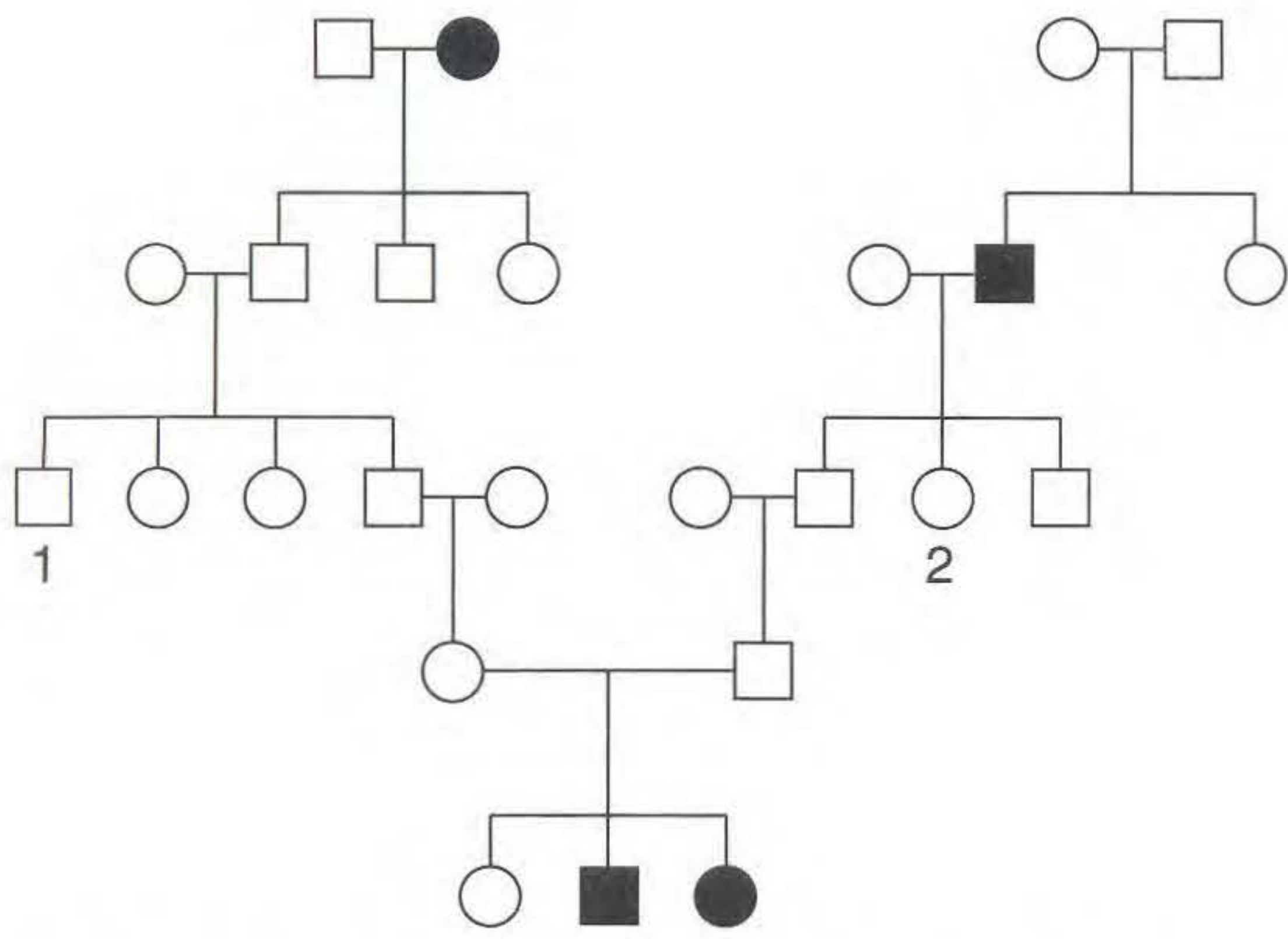


- Como esse distúrbio é herdado? Cite os motivos para sua resposta.
- Cite os genótipos de tantos indivíduos quanto possível no heredograma. (Invente seus próprios símbolos alélicos definidos.)
- Considere as quatro crianças não-afetadas dos genitores III-4 e III-5. Em todas as proles com quatro filhos de genitores desses genótipos, que proporção é esperada contendo todos os filhos não-afetados?

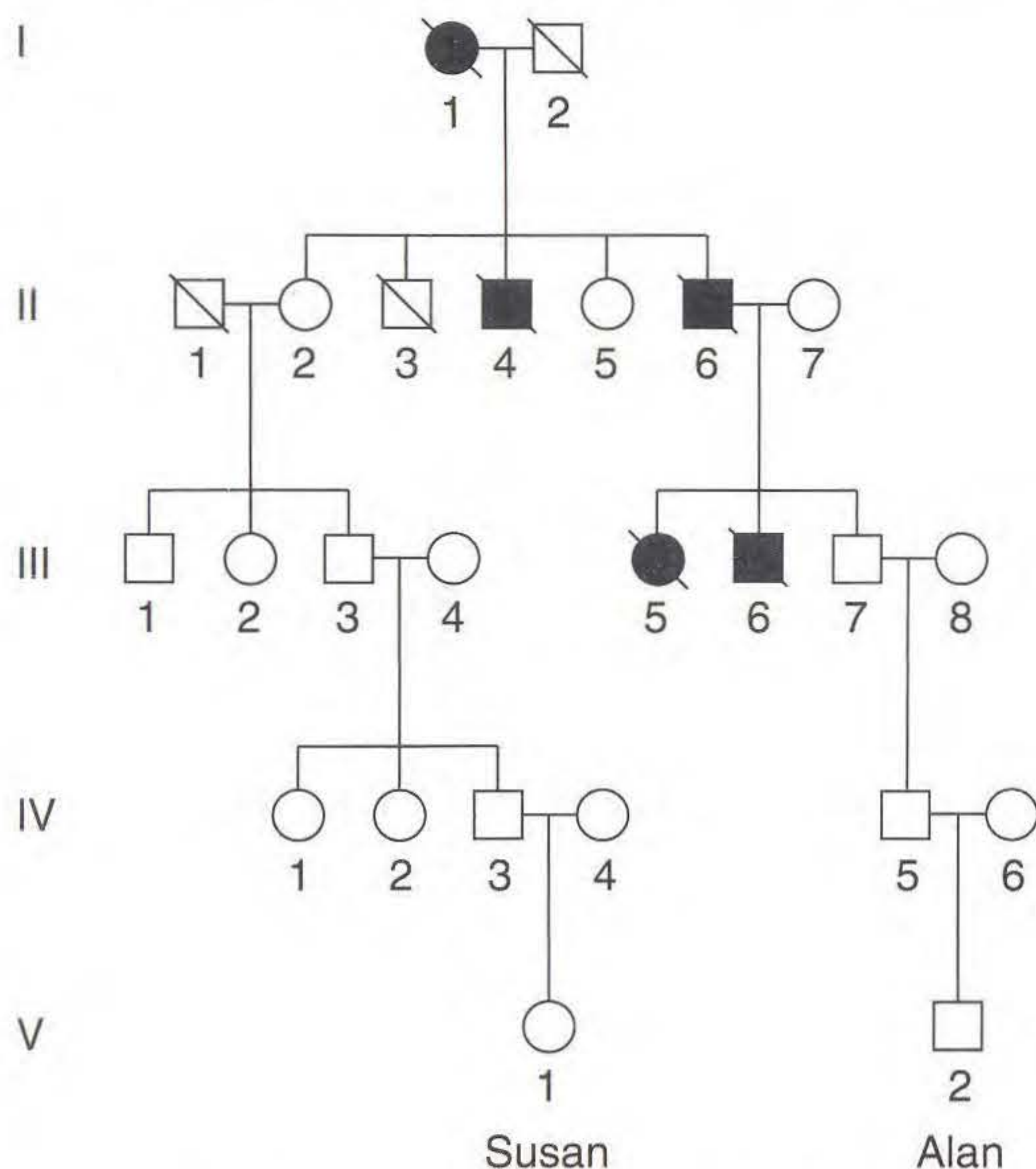
38. São mostrados quatro heredogramas humanos na ilustração seguinte. Os símbolos escuros representam um fenótipo anormal herdado de modo mendeliano simples.



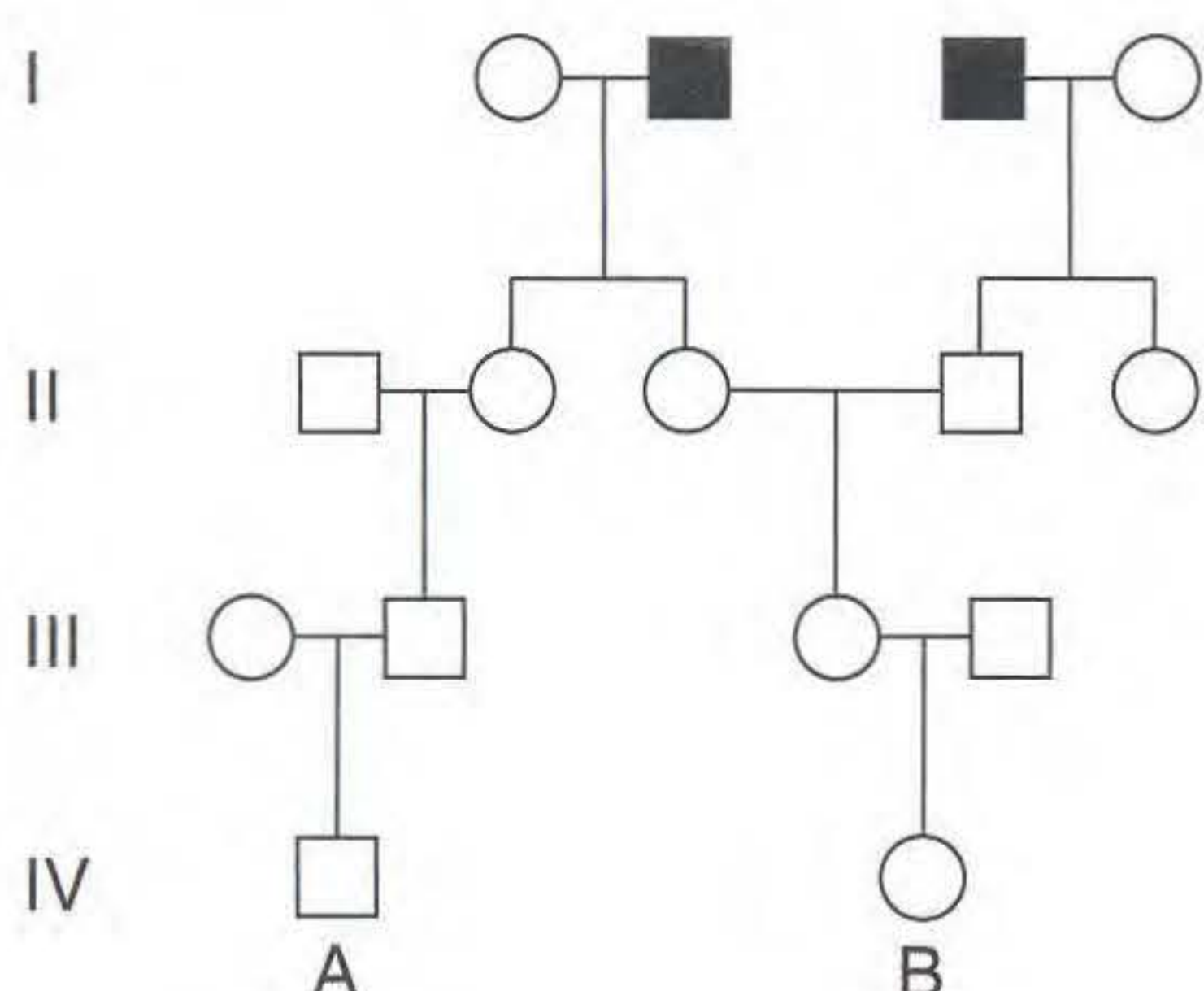
- Para cada heredograma, diga se a condição anormal é dominante ou recessiva. Tente citar a lógica subjacente a sua resposta.
 - Para cada heredograma, descreva os genótipos de tantas pessoas quanto possível.
39. Doença de Tay-Sachs (idiotia infantil amaurótica) é uma doença humana rara na qual substâncias tóxicas se acumulam nas células nervosas. O alelo recessivo responsável pela doença é herdado de modo mendeliano simples. Por motivos desconhecidos, o alelo é mais comum em populações de judeus Ashkenazi do Leste da Europa. Uma mulher está planejando casar com seu primo em primeiro grau, mas o casal descobre que a irmã de seu avô compartilhado morreu na lactância de doença de Tay-Sachs.
- Desenhe as partes relevantes do heredograma e mostre todos os genótipos o mais completamente possível.
 - Qual a probabilidade de que o primeiro filho dos primos tenha doença de Tay-Sachs, supondo que todas as pessoas que entram para a família sejam homozigotas normais?
40. O heredograma a seguir foi obtido de uma rara doença renal.



- a. Deduza a herança dessa condição, citando seus motivos.
 b. Se as pessoas 1 e 2 se casam, qual a probabilidade de que seu primeiro filho tenha a doença renal?
41. Esse heredograma é da doença de Huntington, um distúrbio de manifestação tardia do sistema nervoso. As barras indicam membros falecidos da família.

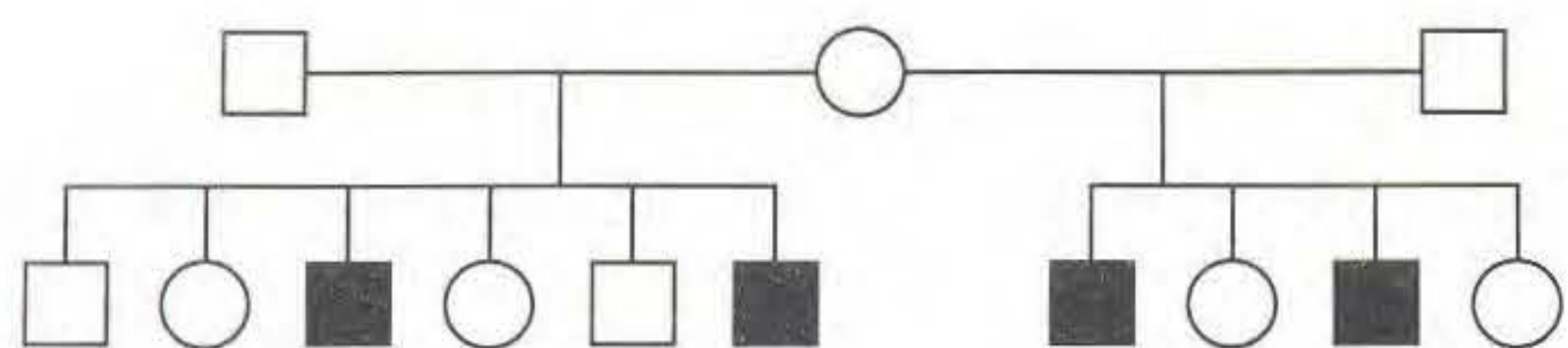


- a. Esse heredograma é compatível com o modo de herança para doença de Huntington mencionado neste capítulo?
 b. Considere duas crianças neonatas nos dois ramos do heredograma, Susan no lado esquerdo e Alan no lado direito. Estude o gráfico na Figura 2.32 e forme uma opinião sobre a probabilidade de que eles desenvolverão doença de Huntington. Suponha para fins de discussão que os genitores tiveram filhos aos 25 anos.
42. Considere o heredograma seguinte de uma rara doença autossômica recessiva, PKU.



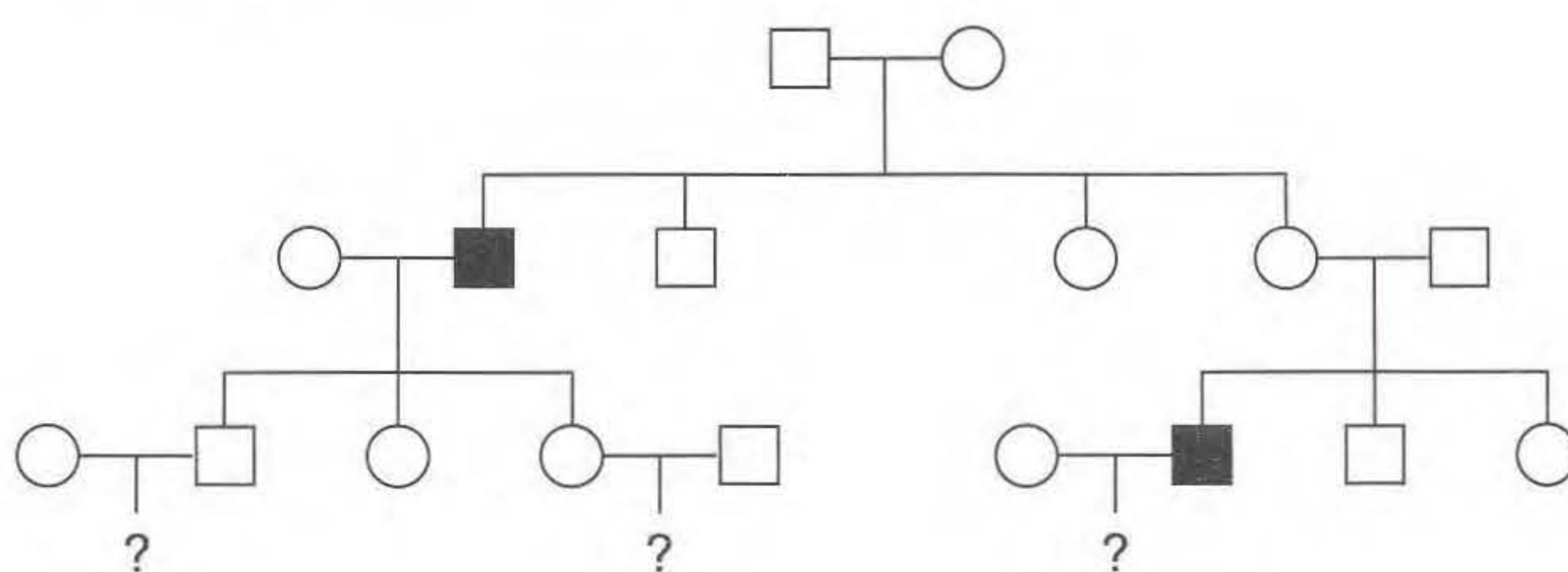
- a. Cite os genótipos de tantas pessoas da família quanto possível.
 b. Se as pessoas A e B se casam, qual a probabilidade de que seu primeiro filho tenha PKU?
 c. Se seu primeiro filho é normal, qual a probabilidade de que seu segundo filho tenha PKU?
 d. Se seu primeiro filho tem a doença, qual a probabilidade de que seu segundo filho não seja afetado?
 (Suponha que todas as pessoas que entram para a família não tenham o alelo anormal.)
43. Um homem não tem lobos na orelha, enquanto sua esposa tem. Seu primeiro filho, um menino, não tem lobos na orelha.
 a. Se a diferença fenotípica é suposta como sendo devida a dois alelos de um único gene, é possível que o gene seja ligado ao X?
 b. É possível decidir se a falta de lobos na orelha é dominante ou recessiva?
44. Um alelo recessivo raro herdado de modo mendeliano causa a doença fibrose cística. Um homem fenotipicamente normal cujo pai teve fibrose cística casa-se com uma mulher fenotipicamente normal de fora da família, e o casal pensa em ter filhos.
 a. Desenhe o heredograma como descrito.
 b. Se a frequência na população de heterozigotos para fibrose cística é de 1 em 50, qual a chance de que o primeiro filho do casal tenha fibrose cística?
 c. Se o primeiro filho tem fibrose cística, qual a probabilidade de que o segundo filho seja normal?
45. O alelo c causa albinismo em camundongos (C faz com que os camundongos sejam pretos). O cruzamento $C/c \times c/c$ produz uma prole de 10. Qual a probabilidade de que todos sejam pretos?
46. O alelo recessivo s faz com que as drosófilas tenham asas pequenas, e o alelo s^1 causa asas normais. Esse gene é conhecido como ligado ao X. Se um macho de asas pequenas é cruzado com uma fêmea homozigota tipo selvagem, que proporção de moscas normais para asas pequenas pode ser esperada em cada sexo na F_1 ? Se as moscas de F_1 forem entrecruzadas, que proporções da prole F_2 são esperadas? Que proporções de prole são previstas se as fêmeas de F_1 forem retrocruzadas com seu pai?
47. Um alelo dominante ligado ao X causa hipofosfatemia em humanos. Um homem com hipofosfatemia casa-se com uma mulher normal. Que proporção de seus filhos terá hipofosfatemia?
48. A distrofia muscular de Duchenne é ligada ao sexo e geralmente afeta apenas homens. As vítimas da doença tornam-se progressivamente mais fracas, começando cedo na vida.
 a. Qual a probabilidade de que uma mulher cujo irmão tem a doença de Duchenne tenha um filho afetado?
 b. Se o irmão de sua mãe (seu tio) teve doença de Duchenne, qual a probabilidade de que você tenha recebido o alelo?
 c. Se o irmão de seu pai teve a doença, qual a probabilidade de que você tenha recebido o alelo?

49. Um homem e uma mulher recém-casados descobrem que cada um teve um filho com alcaptonúria, também conhecida como “doença da urina preta”, uma doença rara causada por um alelo autossômico recessivo de um único gene. Eles estão prestes a ter seu primeiro filho. Qual a probabilidade de que seu filho tenha alcaptonúria?
50. O heredograma seguinte é de uma anomalia dentária, amelogenese imperfeita.



- a. Qual o modo de herança que melhor explica a transmissão da característica?
- b. Escreva os genótipos de todos os membros da família de acordo com sua hipótese.
51. Um casal prestes a se casar descobre pelos estudos de suas histórias familiares que, em *ambas* as famílias, seus avós não-afetados tinham irmãos com fibrose cística (uma rara doença autossômica recessiva).
- a. Se o casal se casa e tem um filho, qual a probabilidade de que a criança tenha fibrose cística?
- b. Se eles tiverem quatro filhos, qual a chance de que os filhos tenham exatamente a proporção mendeliana de 3:1 de normais: fibrose cística?
- c. Se seu primeiro filho tiver fibrose cística, qual a probabilidade de que seus próximos três filhos sejam normais?
52. Um alelo recessivo ligado ao sexo *c* produz um daltonismo em humanos. Uma mulher normal cujo pai era daltônico casa-se com um homem daltônico.
- a. Que genótipos são possíveis para a mãe do homem daltônico?
- b. Quais as chances de que o primeiro filho desse casamento seja um menino daltônico?
- c. Das meninas geradas por esses pais, que proporção pode ser esperada sendo daltônica?
- d. De todos os filhos (sexo não-especificado) desses pais, que proporção pode ser esperada como tendo visão de cores normal?
53. Os gatos caseiros masculinos são pretos ou laranja; as fêmeas são pretas, laranja ou calico.
- a. Se esses fenótipos de cor de pelagem são controlados por um gene ligado ao sexo, como essas observações podem ser explicadas?
- b. Usando símbolos apropriados, determine os fenótipos esperados na prole de um cruzamento entre uma fêmea laranja e um macho preto.
- c. Metade das fêmeas produzidas por um certo tipo de cruzamento é calico e metade é preta; metade dos machos é laranja e metade é preto. De que cores são os machos e fêmeas parentais nesse tipo de cruzamento?
- d. Outro tipo de cruzamento produz uma prole nas seguintes proporções: um quarto de machos laranja, um quarto de fêmeas laranja, um quarto de machos pretos e um quarto de fêmeas calico. De que cores são os machos e fêmeas genitores nesse tipo de cruzamento?

54. O heredograma a seguir é de uma certa doença rara que é incapacitante mas não fatal.



- a. Determine o modo mais provável de herança dessa doença.
- b. Escreva o genótipo de cada membro familiar de acordo com seu modo de herança proposto.
- c. Se você fosse o médico dessa família, como você informaria os três casais na terceira geração sobre a probabilidade de ter um filho afetado?
55. Em milho, um alelo causa um endosperma adocicado, enquanto *S* causa amiláceo. Que genótipos de endospermas resultam de cada um dos seguintes cruzamentos?
- a. fêmea *s/s* × macho *S/S*
- b. fêmea *S/S* × macho *s/s*
- c. fêmea *S/s* × macho *S/s*
56. Uma geneticista de plantas tem duas linhagens puras, uma com pétalas púrpura e uma com azul. Ela supõe que a diferença fenotípica é devida a dois alelos de um gene. Para testar essa idéia, ela procura uma proporção de 3:1 na *F*₂. Ela cruza as linhagens e encontra que toda a prole *F*₁ é púrpura. As plantas *F*₁ são autofecundadas e são obtidas 400 plantas de *F*₂, sendo 320 púrpura e 80 azuis. Esse resultado ajusta-se bem à sua hipótese? Caso não, sugira por quê.
57. O avô de um homem tem galactosemia, uma doença rara autossômica recessiva causada pela incapacidade em processar galactose, levando à disfunção de músculos, nervos e rins. O homem casa-se com uma mulher cuja irmã tem galactosemia. A mulher agora está grávida de seu primeiro filho.
- a. Desenhe o heredograma descrito.
- b. Qual a probabilidade de que essa criança tenha galactosemia?
- c. Se o primeiro filho tem galactosemia, qual a probabilidade de que o segundo filho a tenha?



PROBLEMAS DESAFIADORES

58. Um geneticista trabalhando com ervilhas tem uma única planta monoíbrida *Y/y* (amarela) e, com a autofecundação dessa planta, quer produzir uma planta com genótipo *y/y* para usar como testador. Quantas plantas da prole precisam ser cultivadas para se estar 95% seguro de obter pelo menos uma na amostra?
59. Um curioso polimorfismo em populações humanas é relacionado à habilidade de dobrar os lados da língua para fazer um U. Algumas pessoas podem fazer isso, e outras simplesmente não. Assim, isso é um exemplo de dimorfismo. Seu significado é um completo mistério. Em

uma família, um menino é incapaz de enrolar a língua, mas, para sua grande contrariedade, sua irmã pode fazê-lo. Além disso, ambos os pais enrolam, bem como seus avós, um tio paterno e uma tia paterna. Uma tia paterna, um tio paterno e um tio materno não enrolam a língua.

a. Desenhe o heredograma para essa família, definindo claramente seus símbolos, e deduza os genótipos de tantos membros quanto possível.

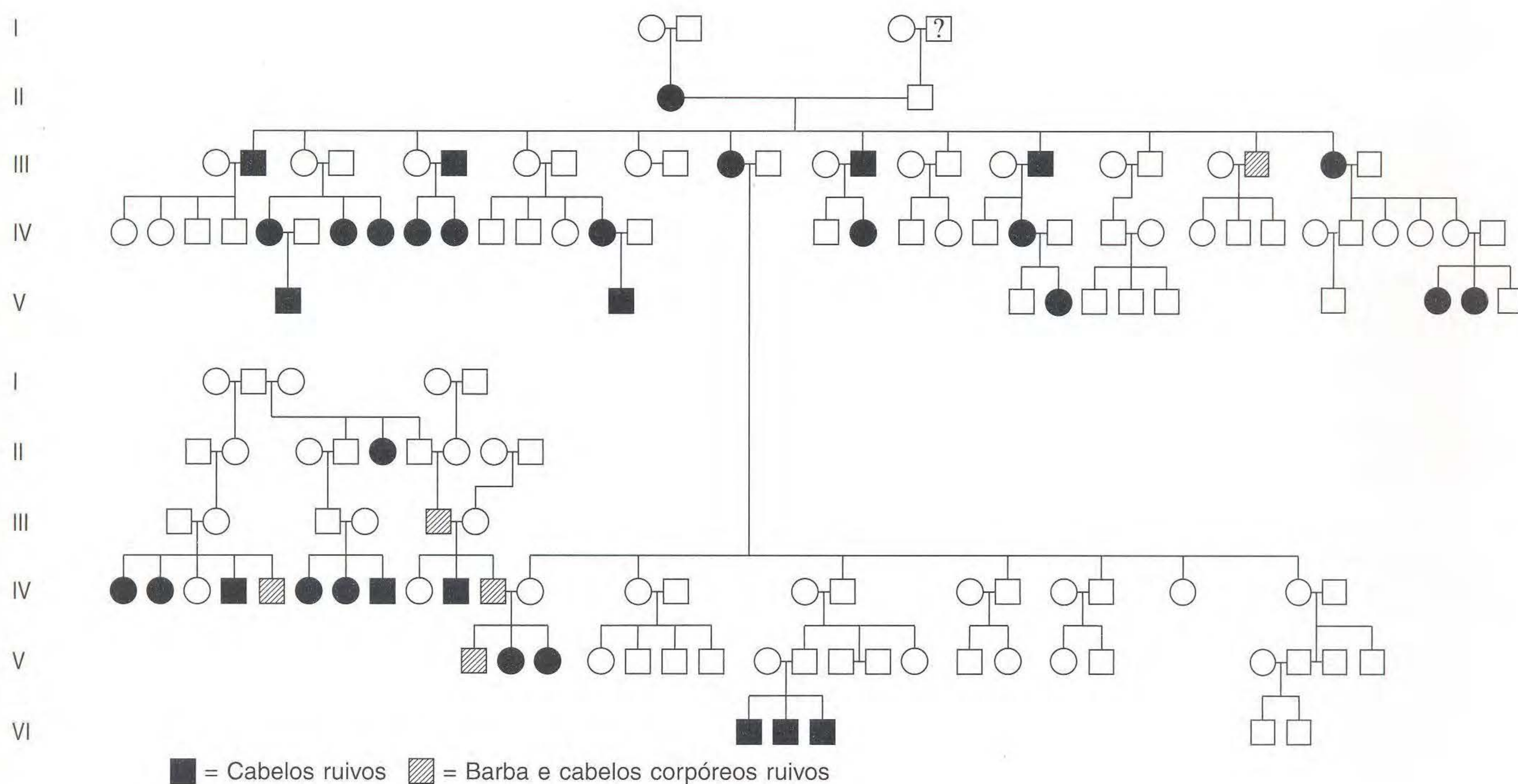
b. O heredograma que você desenhou é típico da herança de enrolar a língua, e levou os geneticistas a avaliar o mecanismo de herança que, sem dúvida, você deduziu. Entretanto, em um estudo de 33 pares de gêmeos idênticos, ambos os membros de 18 pares podiam enrolar, nenhum dos membros de 8 pares podia enrolar e um dos gêmeos em 7 pares podia enrolar, mas o outro não.

Como os gêmeos idênticos são derivados da divisão de um zigoto em dois embriões, os membros do par devem ser geneticamente idênticos. Como a existência dos sete pares discordantes pode ser conciliada com sua explicação genética do heredograma?

60. Cabelos ruivos ocorrem em famílias, e a ilustração a seguir mostra um grande heredograma para cabelos ruivos. (Heredograma de W.R. Singleton and B. Ellis, *Journal of Heredity* 55, 1964, 261.)

a. O padrão de herança nesse heredograma sugere que cabelos ruivos podem ser causados por um alelo dominante ou um recessivo que é herdado de modo mendeliano simples?

b. Você acha que o alelo de cabelos ruivos é comum ou raro na população como um todo?



61. Quando muitas famílias são testadas quanto à habilidade em sentir o gosto de feniltiocarbamida, as reproduções foram agrupadas em três tipos e a prole foi totalizada, com os resultados mostrados a seguir:

Genitores	Número de famílias	Crianças	
		Sensíveis	Insensíveis
Sensíveis × sensíveis	425	929	130
Sensíveis × insensíveis	289	483	278
Insensíveis × insensíveis	86	5	218

Com a suposição de que sentir o gosto de PTC é dominante (P) e não sentir é recessivo (p), como as proporções da prole em cada um desses três tipos de casamentos podem ser explicadas?

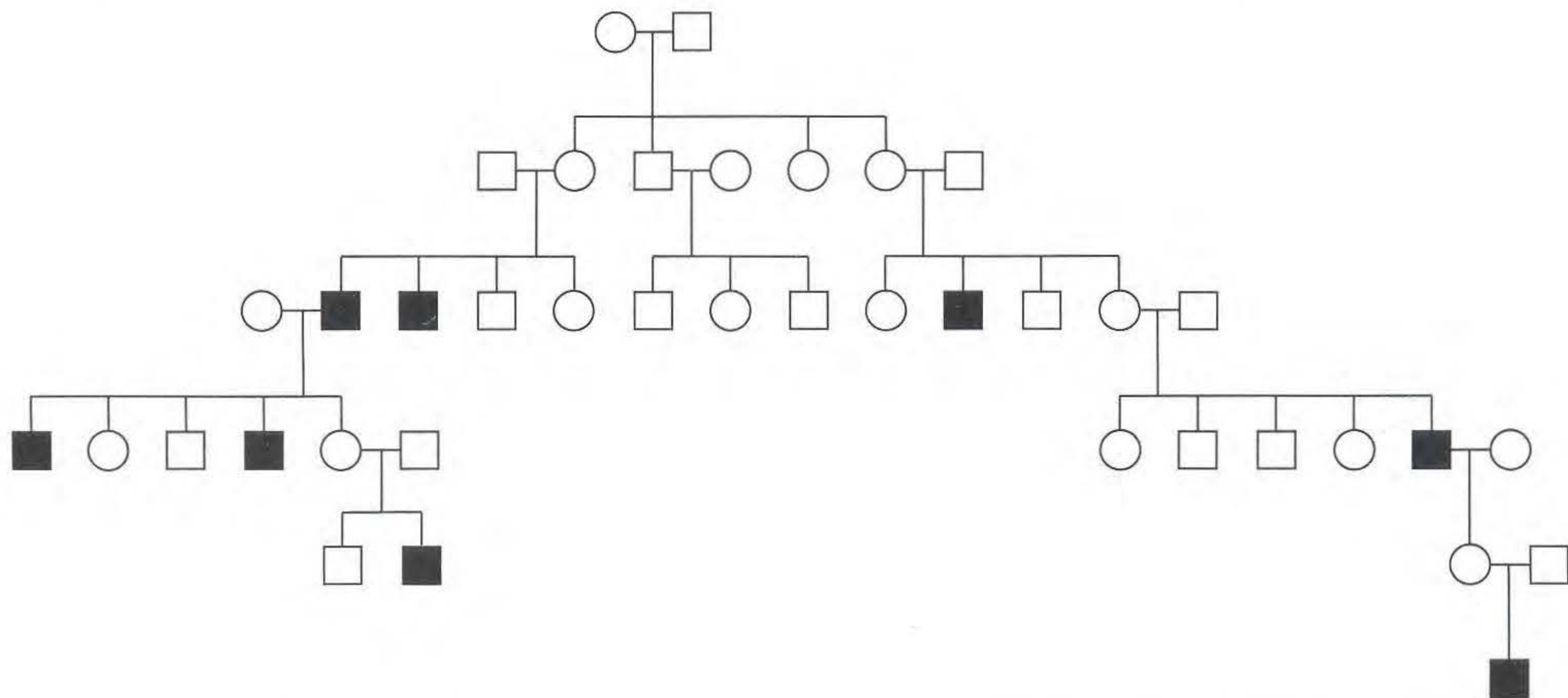
62. Uma condição conhecida como *ictiosis histrix gravior* apareceu em um menino no início do século XVIII. Sua pele tornou-se muito grossa e formou espinhos que se desprendiam a intervalos. Quando ele cresceu, esse “homem porcupino” casou-se e teve seis filhos, todos os quais tinham essa condição, e várias filhas, todas normais. Por quatro gerações, essa condição foi transmitida de pai para filho. Por essa evidência, o que você pode dizer sobre a localização do gene?

63. O tipo selvagem (W) da mariposa *Abraxas* tem grandes manchas em suas asas, mas a forma laticolor (L) dessa espécie tem manchas pequenas. Foram feitos cruzamentos entre as linhagens que diferem nessa característica, com os seguintes resultados:

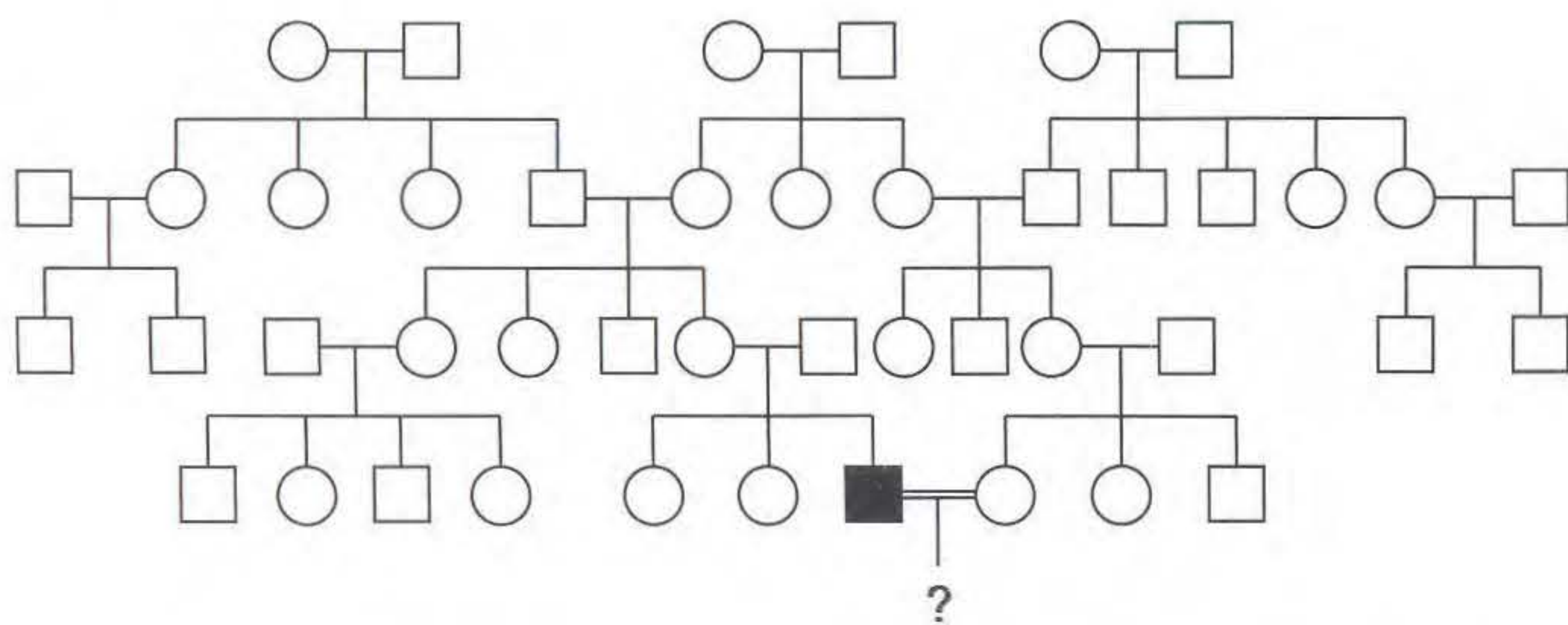
Prole	Genitores		Cruzamento	
	♀	♂	F ₁	F ₂
1	L	W	♀ W	♀ $\frac{1}{2}$ L, $\frac{1}{2}$ W
			♂ W	♂ W
2	W	L	♀ L	♀ $\frac{1}{2}$ W, $\frac{1}{2}$ L
			♂ W	♂ $\frac{1}{2}$ W, $\frac{1}{2}$ L

Forneça uma explicação clara dos resultados nesses dois cruzamentos, mostrando os genótipos de todas as mariposas.

64. O heredograma a seguir mostra a herança de uma rara doença humana. O padrão é mais bem explicado como sendo causado por um alelo recessivo ligado ao X ou por um alelo autossômico dominante com expressão limitada aos homens? (Heredograma modificado de J.F. Crow, *Genetic Notes*, 6th ed. Copyright 1967 by Burgess Publishing Co., Minneapolis.)

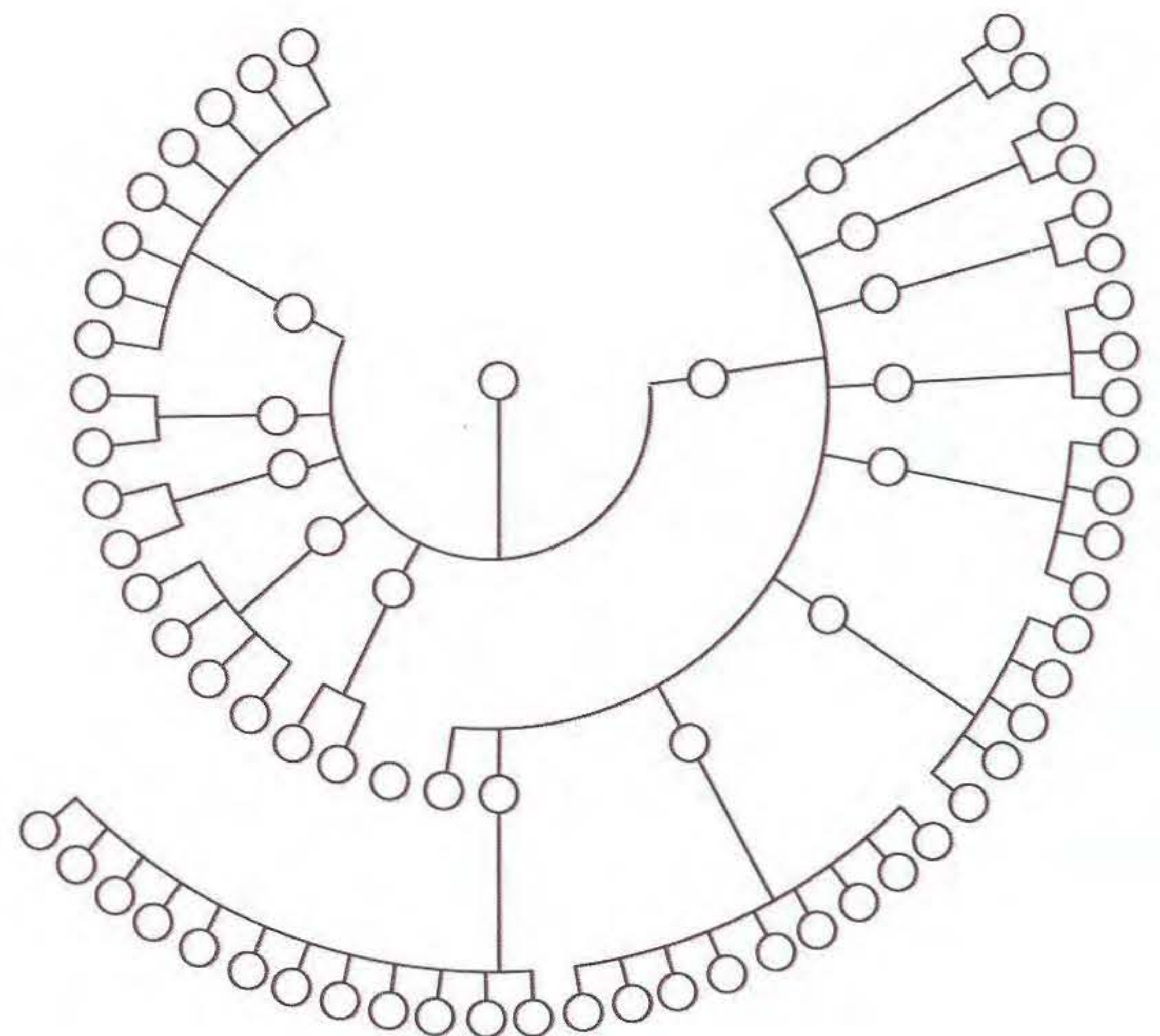


65. Um certo tipo de surdez em humanos é herdado como uma característica recessiva ligada ao X. Um homem que sofre desse tipo de surdez casou-se com uma mulher, e estão esperando um filho. Eles descobrem que são distantemente aparentados. Parte do heredograma é mostrado aqui.



Como você alertaria os parentes sobre a probabilidade de seus filhos terem um filho surdo, uma filha surda, um menino normal ou uma menina normal? Certifique-se de todas as suas suposições.

66. O heredograma acompanhante mostra um padrão de herança muito incomum que de fato existe. Toda a prole é mostrada, mas os pais em cada reprodução foram omitidos para chamar a atenção para o padrão marcante.
- a. Resumidamente, diga o que é incomum sobre esse heredograma.
- b. O padrão pode ser explicado por herança mendeliana?



EXPLORANDO GENOMAS Um Tutorial de Bioinformática Baseado na Web

OMIM e a Doença de Huntington

O Online Mendelian Inheritance in Man (OMIM) coleta dados sobre a genética de genes humanos. No tutorial de Genomics

em www.whfreeman.cm/iga9e, você aprenderá como pesquisar o OMIM sobre dados de locus gênico e padrão de herança para condições tais como a doença de Huntington.

Distribuição Independente de Genes

3



A Revolução Verde na agricultura é propiciada pelo amplo plantio de linhagens superiores de lavouras (tais como arroz, mostrado aqui) que combinam características genéticas benéficas. [Jorgen Schytte/Peter Arnold.]

Este capítulo é sobre os princípios que funcionam quando dois ou mais casos de herança monogênica são analisados simultaneamente. Em nenhuma outra parte esses princípios foram mais importantes do que no cruzamento de animais e na agricultura. Por exemplo, entre os anos de 1960 e 2000, a produção mundial de alimentos dobrou. Esse aumento foi chamado de Revolução Verde. O que possibilitou essa Revolução Verde? Em parte, a Revolução Verde foi devida a práticas agrícolas melhoradas, porém mais importante foi o desenvolvimento de genótipos superiores pelos geneticistas de plantas. Esses criadores estão constantemente à procura de uma ocorrência casual de mutações monogênicas que aumentam significativamente a produtividade ou o valor nutritivo. Entretanto, tais mutações surgem em linhagens diferentes em diferentes partes do mundo. Por exemplo, em arroz, um dos principais cultivos de alimentos do mundo, as mutações seguintes foram cruciais na Revolução Verde:

Perguntas Fundamentais

- Que proporções de prole são produzidas por diíbridos, triíbridos e assim por diante, se os genes estão em pares de cromossomos diferentes?
- Como o comportamento cromossômico na meiose explica essas proporções?
- A herança mendeliana pode explicar a variação contínua?
- Quais os padrões de herança dos genes de organelas?

Tópicos

- 3.1 Lei de Mendel da distribuição independente
- 3.2 Trabalhando com a distribuição independente
- 3.3 A base cromossômica da distribuição independente
- 3.4 Herança poligênica
- 3.5 Genes de organelas: herança independente do núcleo

sd1. Esse alelo recessivo resulta em baixa estatura, tornando a planta mais resistente à “derrubada” das plantações pelo vento ou pela chuva. Também aumenta a quantidade relativa de energia da planta que existe na semente, a parte comestível.

se1. Esse alelo recessivo altera os requisitos da planta por luz do dia, permitindo que seja cultivada em latitudes diferentes.

Xa4. Esse alelo dominante confere resistência à doença da praga bacteriana.

bph2. Esse alelo confere resistência a insetos marrons saltadores.

Snb1. Esse alelo confere tolerância a submersões de plantas após intensas chuvas.

Fazer um genótipo verdadeiramente superior, combinando tais alelos em uma linhagem, é claramente desejável. Para obter tal combinação, as linhagens mutantes devem ser entrecruzadas duas de cada vez. Por exemplo, um geneticista de plantas pode começar cruzando *sd1* e *Xa4*. A prole F_1 desse cruzamento levaria ambas as mutações, mas no estado heterozigoto. Entretanto, a maioria dos agricultores usam linhagens puras, que podem ser eficientemente propagadas e distribuídas aos fazendeiros. Para obter uma linhagem pura duplamente mutante *sd1/sd1* · *Xa4/Xa4*, a F_1 teria que ser mais cruzada para permitir que os alelos se “distribuam” na combinação desejável. Alguns produtos de tais cruzamentos são mostrados na Figura 3.1. Que princípios são relevantes aqui? Depende muito de os dois genes estarem no mesmo par de cromossomos ou em pares diferentes de cromossomos. Neste último caso, os pares de cromossomos atuam independentemente na meiose, e os alelos de dois pares heterozigotos são ditos apresentando **distribuição independente**.

Este capítulo explica como podemos reconhecer a distribuição independente e como o princípio de distribuição independente pode ser usado na construção de uma linhagem, tanto na agricultura quanto na pesquisa genética básica. (O Capítulo 4 cobre o princípio análogo aplicável a pares de genes heterozigotos no *mesmo* par cromossômico.)

Os procedimentos analíticos que pertencem à distribuição independente de genes foram primeiro desenvolvidos pelo pai da genética, Gregor Mendel. Assim, novamente, voltaremos a seu trabalho como um exemplo prototípico.

► 3.1 Lei de Mendel da Distribuição Independente

Em grande parte de seu trabalho original com ervilhas, Mendel analisou os descendentes de linhagens puras que diferiam em *duas* características. O seguinte simbolismo geral é usado para representar genótipos que incluem dois genes. Se os dois genes estão em cromossomos diferentes, os pares de genes são separados por um ponto e vírgula — por exemplo, *A/a* ; *B/b*. Se estão no mesmo cromossomo, os alelos em um homólogo são escritos adjacentemente, sem pontuação, sendo separados dos do outro homólogo por uma barra — por exemplo, *AB/ab* ou *Ab/aB*. Não existe um simbolismo aceito para situações nas quais não se sabe se os genes estão no mesmo cromossomo ou em cromossomos diferentes. Para essa situação de posição desconhecida, neste livro usaremos um ponto para separar os genes — por exemplo, *A/a* · *B/b*. Lembre, do Capítulo 2, que um heterozigoto para um *único* gene (tal como *A/a*) às vezes é chamado de um **monoíbrido**. Concordantemente, um **duplo** heterozigoto, tal como *A/a* · *B/b*, às vezes é chamado de um **diíbrido**. Estudando **cruzamentos diíbridos** (*A/a* · *B/b* × *A/a* · *B/b*), Mendel chegou a seu segundo importante princípio da hereditariedade.

O par de características com as quais ele começou a trabalhar eram forma da semente e cor da semente. Nós já seguimos o cruzamento monoíbrido para cor da semente (*Y/y* × *Y/y*), que dá uma proporção de prole de 3 amarelos : 1 verde. Os fenótipos da forma da semente (Figura 3.2) são liso (determinado pelo alelo *R*) e rugoso (determinado pelo alelo *r*). O cruzamento monoíbrido *R/r* × *R/r* dá uma proporção de prole de 3 lisos : 1 rugoso, como esperado (veja Quadro 2.1, Capítulo 2). Para fazer um cruzamento diíbrido, Mendel começou com duas linhagens parentais puras. Uma linhagem



FIG. 3.1 Genótipos superiores de cultivos tais como arroz revolucionaram a agricultura. Essa foto mostra alguns dos genótipos principais usados nos programas de cultivo de arroz. [International Rice Research Institute.]

Fenótipos liso e rugoso

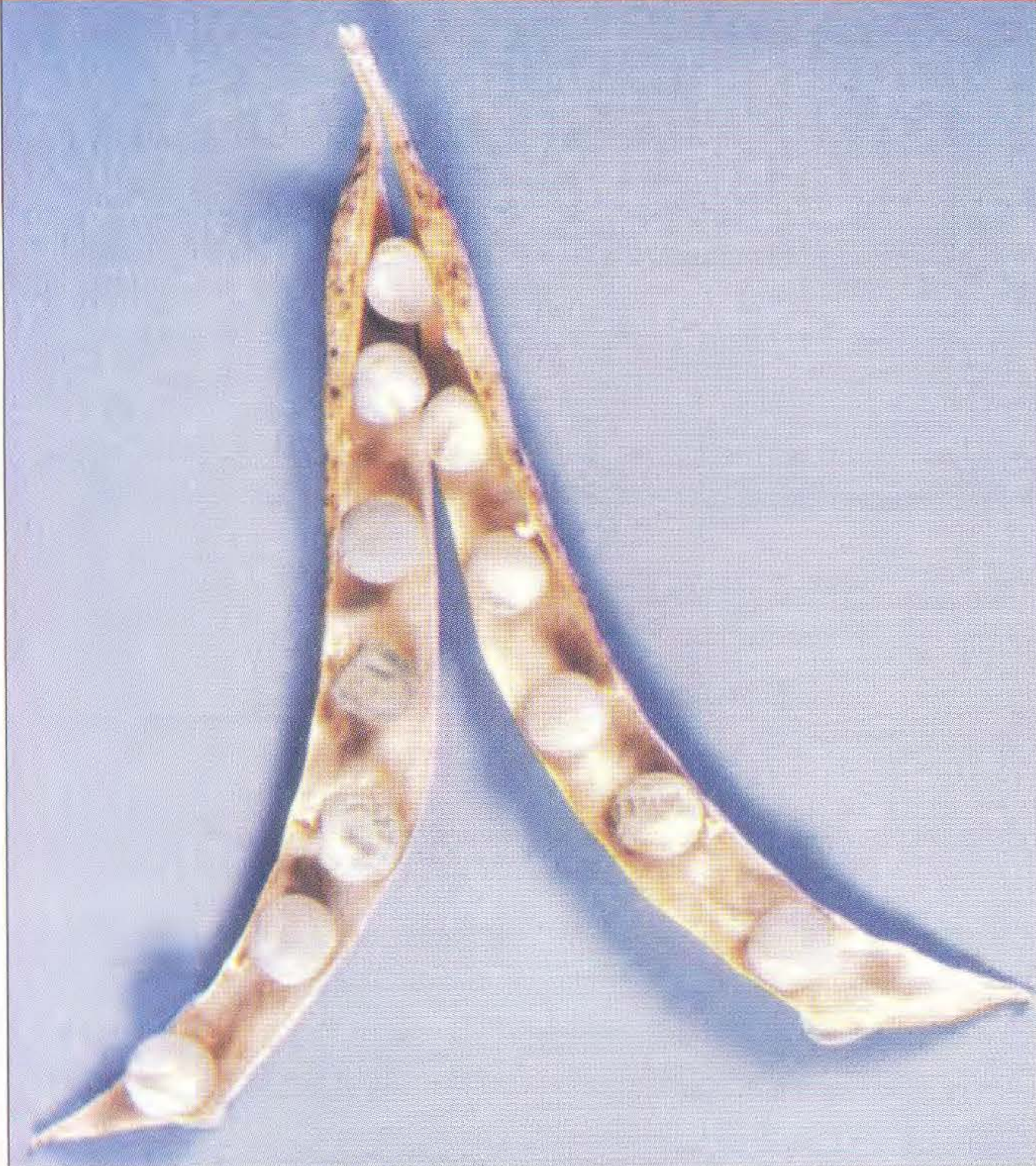


FIG. 3.2 Ervilhas lisas (R/R ou R/r) e rugosas (r/r) estão presentes em uma vagem de plantas heterozigotas autofecundadas (R/r). A proporção fenotípica nessa vagem é exatamente a de 3:1 esperada em média na prole dessa autofecundação. (Estudos moleculares mostraram que o alelo rugoso usado por Mendel é produzido pela inserção de um segmento de DNA móvel no gene; veja Capítulo 14.) [Madan K. Bhattacharyya.]

tinha sementes rugosas, amarelas. Como Mendel não tinha conceito da localização cromossômica dos genes, devemos usar a representação do ponto para escrever o genótipo combinado inicialmente como $r/r \cdot Y/Y$. A outra linhagem tinha sementes lisas, verdes, com o genótipo $R/R \cdot y/y$. Quando essas duas linhagens foram cruzadas, elas devem ter produzido gametas que eram $r \cdot Y$ e $R \cdot y$, respectivamente. Assim, as sementes de F_1 tinham que ser diíbridas, com o genótipo $R/r \cdot Y/y$. Mendel descobriu que as sementes de F_1 eram lisas e amarelas. Esse resultado mostrou que a dominância de R em relação a r e de Y em relação a y não era afetada pela condição do outro par de genes no diíbrido $R/r \cdot Y/y$. Em seguida, Mendel autofecundou os diíbridos de F_1 para obter a geração F_2 .

As sementes de F_2 eram de quatro tipos diferentes nas seguintes proporções:

- $\frac{9}{16}$ lisas, amarelas
- $\frac{3}{16}$ lisas, verdes
- $\frac{3}{16}$ rugosas, amarelas
- $\frac{1}{16}$ rugosa, verde

um resultado que é ilustrado na Figura 3.3 com os números reais obtidos por Mendel. Essa proporção inicialmente inespe-

Programa de cruzamentos de Mendel que produziu uma proporção 9:3:3:1

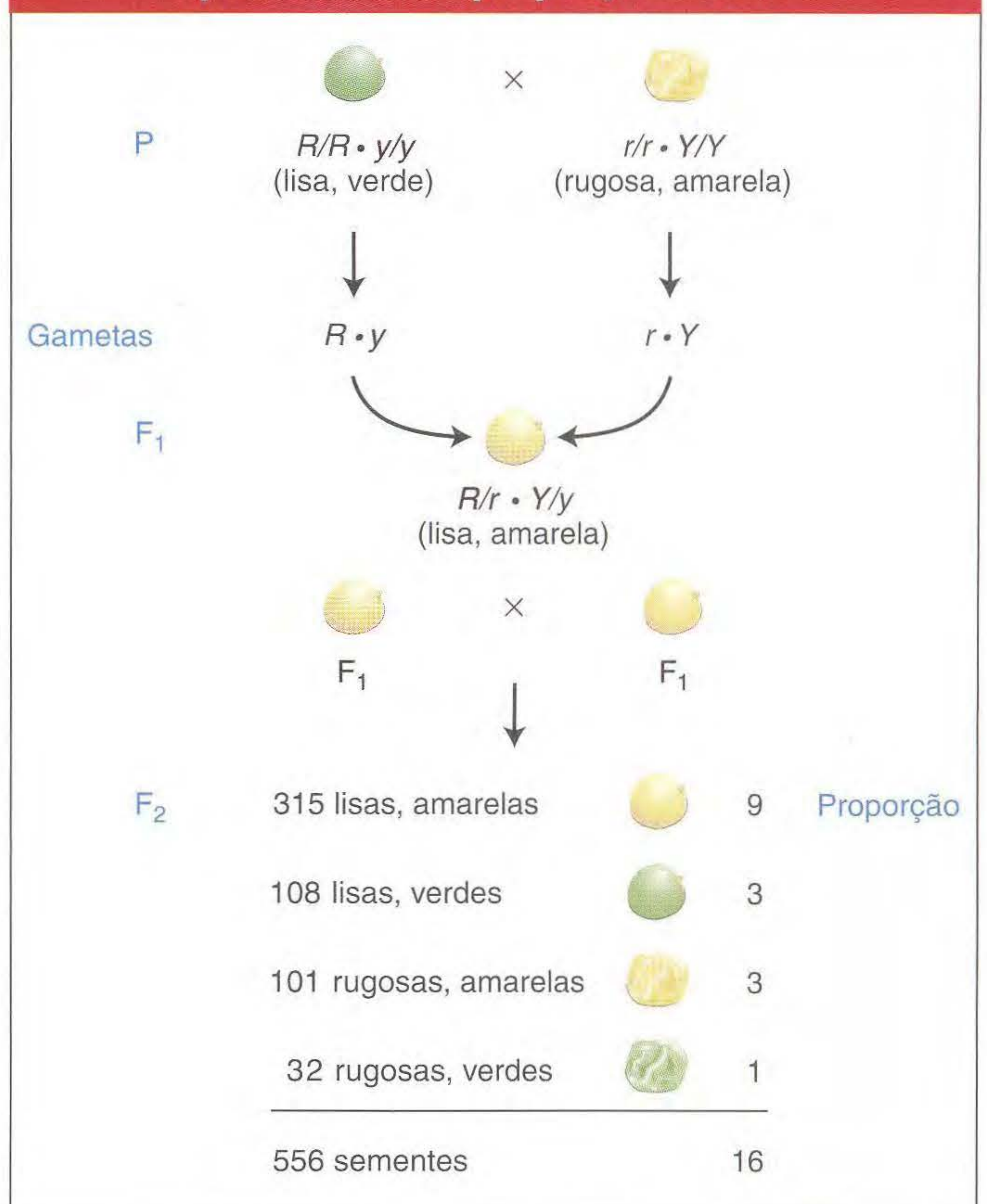
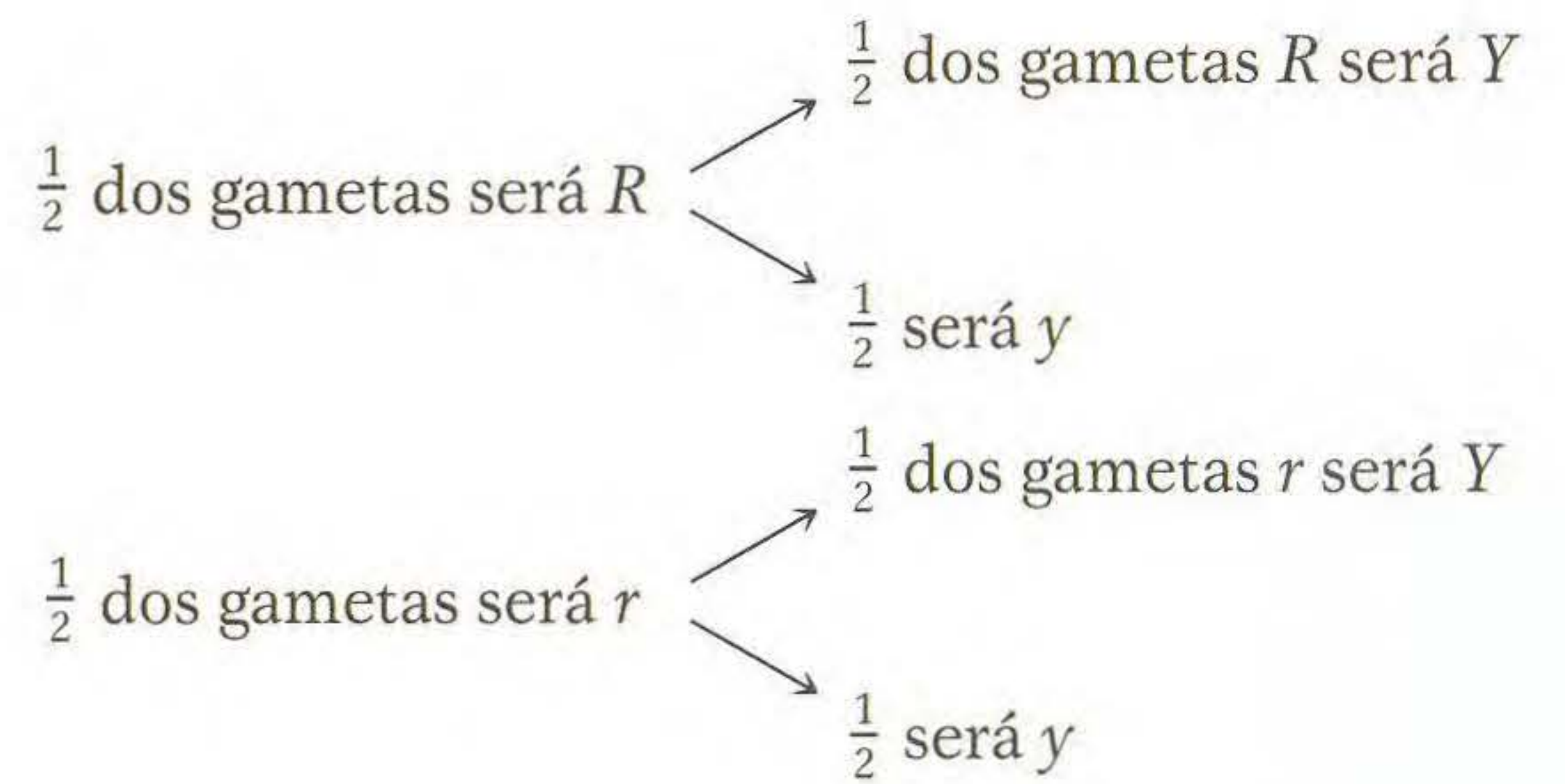
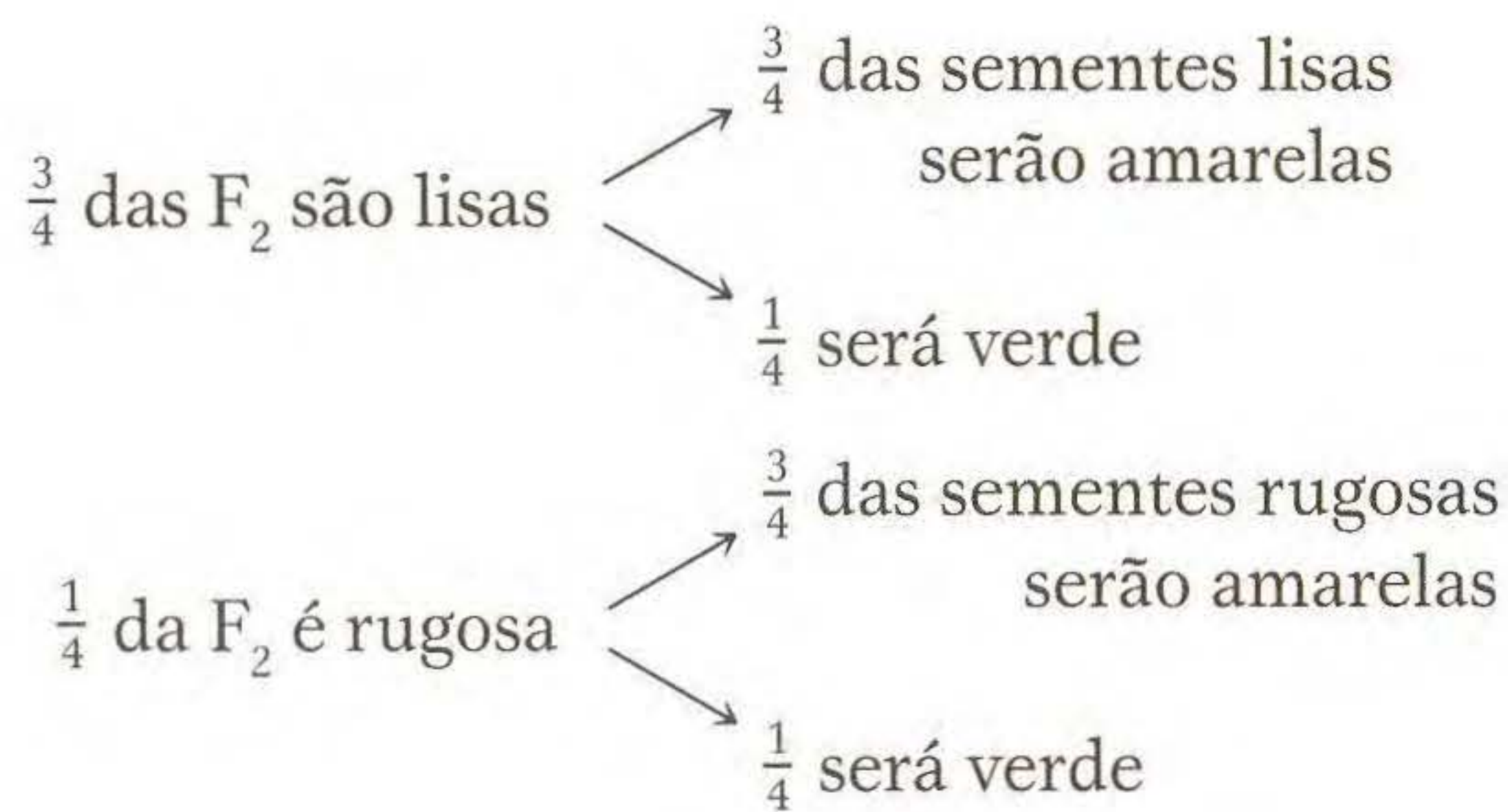


FIG. 3.3 Mendel produziu um diíbrido que, quando autofecundado, produziu uma prole F_2 na proporção 9:3:3:1.

rada de 9:3:3:1 para estas duas características parece muito mais complexa do que as simples proporções de 3:1 dos cruzamentos monoíbridos. Entretanto, a proporção 9:3:3:1 mostrou ser um padrão consistente de herança em ervilhas. Como evidência, Mendel também fez cruzamentos diíbridos que incluíam várias outras combinações de características e encontrou que *todos* os indivíduos diíbridos de F_1 produziram proporções 9:3:3:1 na F_2 . A proporção foi outro padrão de herança que precisou do desenvolvimento de uma nova idéia para explicá-lo.

Primeiro, vamos verificar os números reais obtidos por Mendel na Figura 3.3 para determinar se as proporções monoíbridas 3:1 ainda podiam ser encontradas na F_2 . Com relação à forma da semente, existem 423 sementes lisas (315 + 108) e 133 sementes rugosas (101 + 32). Esse resultado está próximo de uma proporção 3:1. Em seguida, com relação à cor da semente, existem 416 sementes amarelas (315 + 101) e 140 verdes (108 + 32), também muito próximo de uma proporção 3:1. A presença dessas duas proporções 3:1 escondidas na proporção 9:3:3:1 foi, indubitavelmente, a fonte da intuição de que Mendel precisava para explicar a proporção 9:3:3:1, pois ele percebeu que eram simplesmente duas proporções 3:1 combinadas aleatoriamente. Um modo de visualizar a combinação aleatória dessas duas proporções é com um diagrama ramificado, como o seguinte:



As proporções de todos os quatro resultados possíveis são calculadas usando a regra do produto para multiplicar ao longo dos ramos no diagrama. Por exemplo, $\frac{3}{4}$ de $\frac{3}{4}$ é calculado como $\frac{3}{4} \times \frac{3}{4}$, que é igual a $\frac{9}{16}$. Essas multiplicações dão as seguintes quatro proporções:

$$\frac{3}{4} \times \frac{3}{4} = \frac{9}{16} \text{ lisas, amarelas}$$

$$\frac{3}{4} \times \frac{1}{4} = \frac{3}{16} \text{ lisas, verdes}$$

$$\frac{1}{4} \times \frac{3}{4} = \frac{3}{16} \text{ rugosas, amarelas}$$

$$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16} \text{ rugosa, verde}$$

Essas proporções constituem a proporção 9:3:3:1, que estamos tentando explicar. Entretanto, não seria esse exercício apenas um malabarismo de números? O que a combinação das proporções 3:1 significa biologicamente? O modo pelo qual Mendel deu sua explicação de fato indica um mecanismo biológico. No que hoje é conhecido como a **Segunda Lei de Mendel**, ele concluiu que *pares diferentes de genes segregam-se independentemente na formação de gametas*. A consequência é que, para dois pares de genes heterozigotos A/a e B/b , o alelo b termina em um gameta com um alelo a tanto quanto um alelo A , e do mesmo modo para o alelo B . Com percepção, hoje sabemos que, na maior parte, essa “lei” aplica-se a genes em cromossomos diferentes. Genes no mesmo cromossomo geralmente não se segregam independentemente, pois são mantidos juntos no próprio cromossomo. Assim, a versão moderna da Segunda Lei de Mendel é redigida do seguinte modo:

A Segunda Lei de Mendel (o princípio da segregação independente) diz que pares de genes em cromossomos diferentes separam-se independentemente na meiose.

Explicamos a proporção fenotípica 9:3:3:1 como duas proporções fenotípicas 3:1 combinadas aleatoriamente. Podemos também chegar à proporção 9:3:3:1 a partir de uma consideração da frequência de gametas, os produtos meióticos? Consideremos os gametas produzidos pelo diíbrido $R/r ; Y/y$ (o ponto-e-vírgula mostra que agora a idéia é de que os genes estão em cromossomos diferentes). Novamente, usaremos o diagrama ramificado para começar, pois ele ilustra visualmente a independência. Combinando as leis de Mendel de segregação igual e disjunção independente, podemos prever que

A multiplicação ao longo de ramos de acordo com a regra do produto nos dá as proporções de gametas:

$$\frac{1}{4} R ; Y$$

$$\frac{1}{4} R ; y$$

$$\frac{1}{4} r ; Y$$

$$\frac{1}{4} r ; y$$

Essas proporções são um resultado direto da aplicação de duas leis mendelianas: da segregação e da independência. Entretanto, ainda não chegamos à proporção 9:3:3:1. A etapa seguinte é reconhecer que, como Mendel não especificou regras diferentes para a formação de gametas masculinos e femininos, ambos os gametas masculino e feminino apresentarão as mesmas proporções dadas. Os quatro tipos gaméticos femininos serão fertilizados aleatoriamente por quatro tipos gaméticos masculinos para obter a F_2 . O melhor modo gráfico de mostrar os resultados dos cruzamentos é usando uma grade 4×4 , chamada de *quadrado de Punnett*, que é mostrada na Figura 3.4. Já vimos que as grades são úteis em genética para dar uma representação visual dos dados. Sua utilidade está no fato de que suas proporções podem ser obtidas de acordo com as proporções genéticas em consideração. No quadrado de Punnett na Figura 3.4, por exemplo, quatro fileiras e quatro colunas foram desenhadas para corresponder, em tamanho, aos quatro genótipos dos gametas femininos e aos quatro gametas masculinos. Vemos que existem 16 boxes representando as várias fusões gaméticas e que cada box é $1/16$ da área total da grade. De acordo com a regra do produto, cada $1/16$ é um resultado da fertilização de um tipo de ovócito com uma frequência de $\frac{1}{4}$ por um tipo de espermatozóide também na frequência de $\frac{1}{4}$, dando a frequência dessa fusão como $(\frac{1}{4})^2$. Como mostra o quadrado de Punnett, a F_2 contém uma variedade de genótipos, mas existem apenas quatro fenótipos e suas proporções são 9:3:3:1. Assim, vemos que, quando calculamos as frequências da prole diretamente pelas frequências gaméticas, ainda chegamos à proporção 9:3:3:1. Assim, as leis de Mendel explicam não só os fenótipos de F_2 , mas também os genótipos dos gametas e a prole subjacente à proporção fenotípica de F_2 .

Mendel testou seu princípio de distribuição independente de vários modos. O modo mais direto enfocou a proporção gamética 1:1:1:1 supostamente produzida pelo diíbrido de $F_1 R/r ; Y/y$, pois essa proporção vem diretamente de seu

Quadrado de Punnett ilustrando os genótipos subjacentes à proporção 9:3:3:1

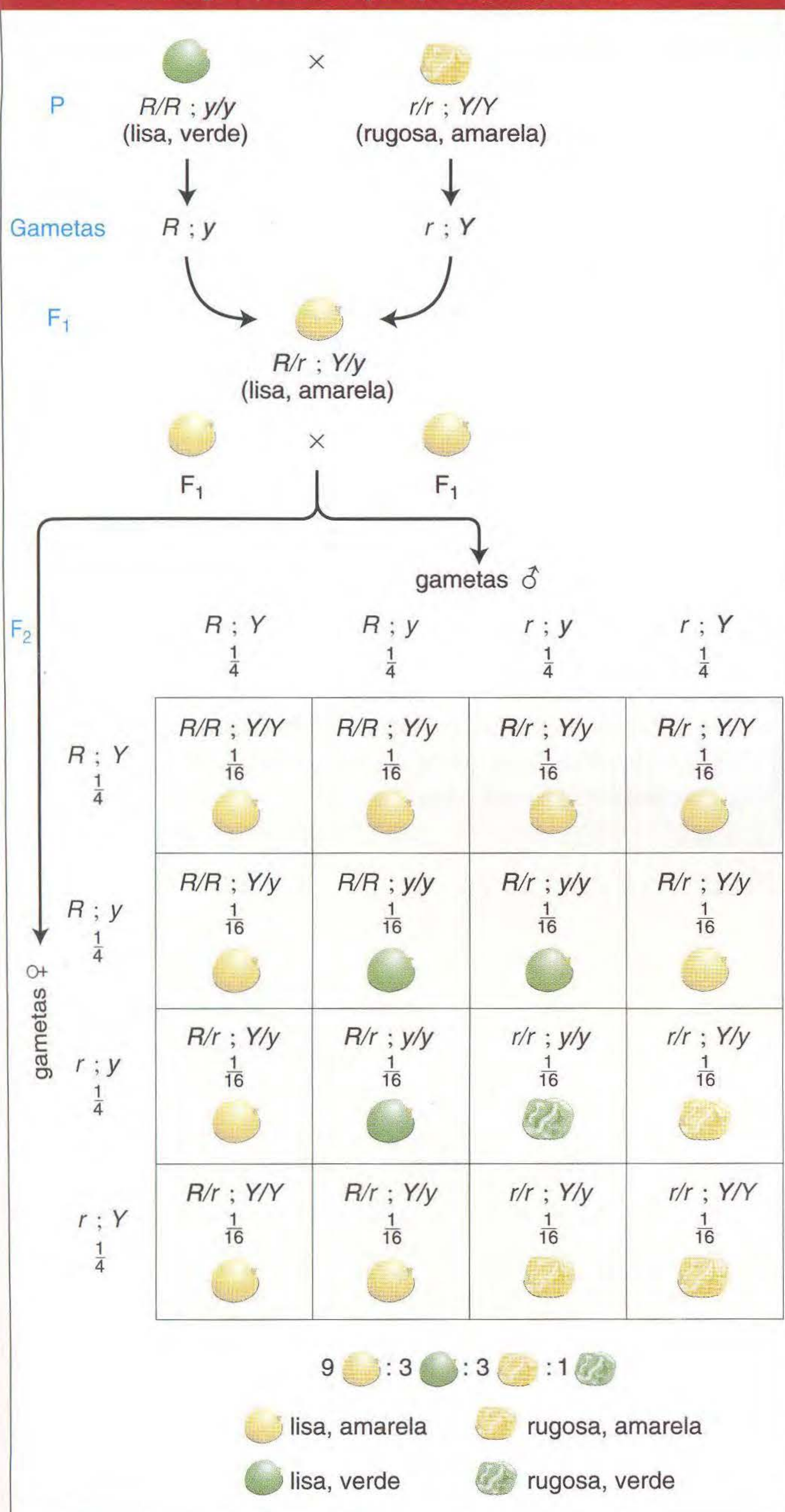


FIG. 3.4 Podemos usar o quadrado de Punnett para prever o resultado de um cruzamento diíbrido. Esse quadrado de Punnett mostra a constituição genotípica e fenotípica prevista da geração F₂ de um cruzamento diíbrido.

princípio de distribuição independente e foi a base biológica da proporção 9:3:3:1 na F₂, como mostrado no quadrado de Punnett. Para verificar a proporção gamética 1:1:1:1, Mendel usou um cruzamento-teste. Ele fez um cruzamento-teste do diíbrido de F₁ com um testador do genótipo $r/r ; y/y$ que só produz gametas com alelos recessivos (genótipo $r ; y$). Ele raciocinou que, se houver de fato uma proporção 1:1:1:1 de $R ; Y, R ; y, r ; Y$ e $r ; y$ gametas, as proporções da prole desse

cruzamento devem corresponder diretamente a proporções gaméticas produzidas pelo diíbrido; em outras palavras,

$$\begin{aligned} &\frac{1}{4} R/r ; Y/y \\ &\frac{1}{4} R/r ; y/y \\ &\frac{1}{4} r/r ; Y/y \\ &\frac{1}{4} r/r ; y/y \end{aligned}$$

Essas proporções foram o resultado que ele obteve, perfeitamente consistentes com suas expectativas. Ele obteve resultados similares para todos os outros cruzamentos diíbridos que fez, e esses testes e outros tipos de testes, todos mostraram que ele de fato criou um modelo robusto para explicar os padrões de herança observados em seus vários cruzamentos de ervilhas.

No início dos anos de 1900, ambas as leis de Mendel foram testadas em um largo espectro de organismos eucarióticos. Os resultados desses testes mostraram que os princípios de Mendel eram em geral aplicáveis. As proporções mendelianas (tais como 3:1, 1:1, 9:3:3:1 e 1:1:1:1) foram amplamente reportadas, sugerindo que a segregação e a distribuição independente são processos hereditários fundamentais encontrados na natureza. As leis de Mendel não são apenas leis sobre ervilhas, mas leis sobre a genética de organismos eucarióticos em geral.

Como um exemplo da aplicabilidade universal do princípio da distribuição independente, podemos examinar sua ação nos haplóides. Se o princípio da segregação igual é válido de modo geral, então devemos ser capazes de observar sua ação nos haplóides, considerando que os haplóides sofrem meiose. De fato, a distribuição independente pode ser observada em um cruzamento do tipo $A ; B \times a ; b$. A fusão de células parentais resulta em um meiócito diplóide transitório que é um diíbrido $A/a ; B/b$, e os produtos aleatoriamente amostrados da meiose (esporos sexuais tais os ascósporos nos fungos) serão

$$\begin{aligned} &\frac{1}{4} A ; B \\ &\frac{1}{4} A ; b \\ &\frac{1}{4} a ; B \\ &\frac{1}{4} a ; b \end{aligned}$$

Assim, vemos a mesma proporção que no cruzamento-teste diíbrido em um organismo diplóide; novamente, a proporção é uma combinação aleatória de duas proporções monoíbridas 1:1 devido à distribuição independente.

Proporções de 1:1:1:1 e 9:3:3:1 são diagnósticas da distribuição independente em um e dois meiócitos diíbridos, respectivamente.

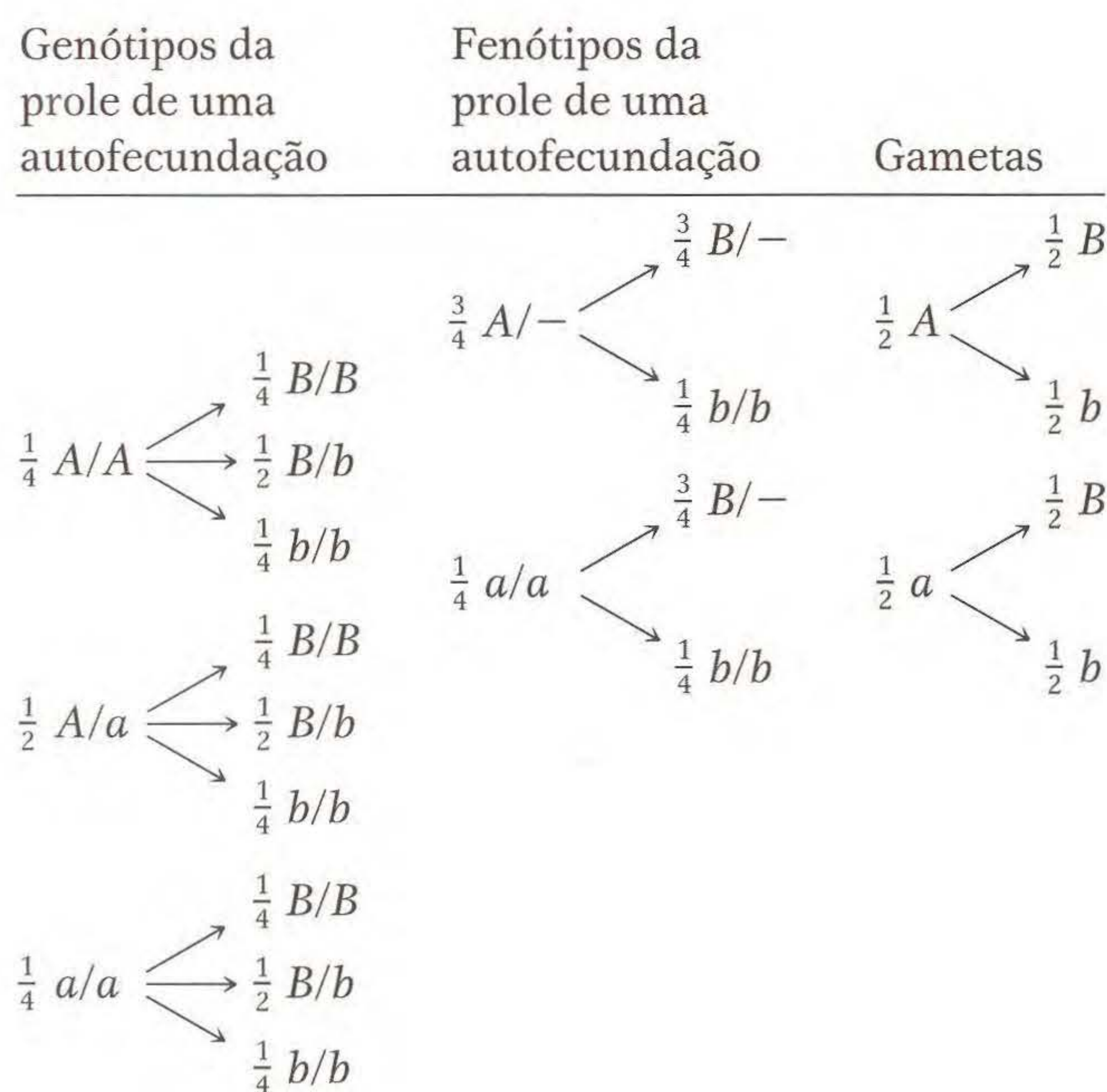
► 3.2 Trabalhando com a Distribuição Independente

Nesta seção, examinaremos vários procedimentos analíticos que são parte das pesquisas genéticas cotidianas, e são todas

baseadas no conceito da distribuição independente. Esses procedimentos lidam com vários aspectos de análise de proporções fenotípicas.

Prevendo as proporções na prole

Como dito no Capítulo 2, a genética pode funcionar em duas direções: (1) prevendo os genótipos dos genitores usando proporções da prole ou (2) prevendo proporções da prole a partir de genitores com genótipo conhecido. Esta última é uma parte importante da genética envolvida com a previsão de tipos de prole que emergem de um cruzamento e calculando suas frequências esperadas, em outras palavras, suas probabilidades. Já examinamos dois métodos para fazer isso: os quadrados de Punnett e os diagramas de ramificação. Os quadrados de Punnett podem ser usados para mostrar os padrões de hereditariedade com base em um par de genes, dois pares de genes, ou mais. Tais grades são bons utensílios gráficos para representar a prole, mas desenhá-los leva tempo. Mesmo o quadrado de Punnett com 16 compartimentos que usamos para analisar um cruzamento diíbrido leva muito tempo para ser desenhado, mas, para um cruzamento triíbrido, existem 2^3 , ou 8, tipos diferentes de gametas, e o quadrado de Punnett tem 64 compartimentos. Os diagramas ramificados (mostrados aqui) são mais fáceis de criar e são adaptáveis a proporções fenotípicas, genotípicas ou gaméticas, como ilustrado para o diíbrido $A/a ; B/b$.



Note, entretanto, que a “árvore” de ramificações para os genótipos é bem difícil de manejar mesmo nesse caso simples, que usa dois pares de genes, pois existem $3^2 = 9$ genótipos. Para três pares de genes, existem 3^3 , ou 27 genótipos possíveis. Para simplificar esse problema, podemos usar um enfoque estatístico, que constitui um terceiro método para calcular as probabilidades (frequências esperadas) de fenótipos específicos ou genótipos vindos de um cruzamento. As duas regras estatísticas necessárias são a **regra dos produtos** (introduzida no Capítulo 2) e a **regra da soma**, que agora consideraremos juntas.

A regra do produto diz que a probabilidade de eventos independentes ocorrerem **ambos** juntos é o produto de suas probabilidades individuais.

Os resultados possíveis no lançamento de dois dados segue a regra do produto porque o resultado de um dado é independente do outro. Como um exemplo, calculemos a probabilidade, p , de obtermos um par de 4. A probabilidade de um 4 em um dado é $1/6$ porque os dados têm seis lados e apenas um lado leva o número 4. Essa probabilidade é escrita do seguinte modo:

$$p(\text{de um } 4) = \frac{1}{6}$$

Portanto, com uso da regra dos produtos, a probabilidade de um 4 aparecer em ambos os dados é $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$, que é escrito

$$p(\text{de dois } 4) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

Agora a regra da soma:

A regra da soma diz que a probabilidade de **um ou outro** de dois eventos mutuamente exclusivos ocorrerem é a soma de suas probabilidades individuais.

(Note que, na regra do produto, o foco é nos resultados A e B . Na regra da soma, o foco é no resultado A' **ou** A'' .)

Os dados também podem ser usados para ilustrar a regra da soma. Nós já calculamos que a probabilidade de dois 4 é $\frac{1}{36}$. Claramente, com o uso do mesmo tipo de cálculo, a probabilidade de dois 5 será a mesma, ou $\frac{1}{36}$. Agora podemos calcular a probabilidade de dois 4 **ou** dois 5. Como esses resultados são mutuamente exclusivos, a regra da soma pode ser usada para nos dizer que a resposta é $\frac{1}{36} + \frac{1}{36}$, que é $\frac{1}{18}$. Essa probabilidade pode ser escrita do seguinte modo:

$$p(\text{dois } 4 \text{ ou dois } 5) = \frac{1}{36} + \frac{1}{36} = \frac{1}{18}$$

Que proporção da prole será de um genótipo específico?

Agora, podemos voltar para um exemplo genético. Suponha que tenhamos duas plantas com os genótipos

$$A/a ; b/b ; C/c ; D/d ; E/e$$

e

$$A/a ; B/b ; C/c ; d/d ; E/e$$

De um cruzamento entre essas plantas, queremos recuperar uma prole de plantas com o genótipo $a/a ; b/b ; c/c ; d/d ; e/e$ (talvez para o fim de agir como linhagem testadora em um cruzamento-teste). Que proporção da prole deve ser esperada tendo esse genótipo? Se for suposto que todos os pares de genes se distribuem independentemente, então podemos fazer facilmente esse cálculo usando a regra do produto. Os cinco pares de genes diferentes são considerados individualmente, como se fossem cinco cruzamentos separados, e então as probabilidades individuais de obter cada genótipo são multiplicadas juntas para chegar à resposta:

De $A/a \times A/a$, um quarto da prole será a/a .

De $b/b \times B/b$, metade da prole será b/b .

De $C/c \times C/c$, um quarto da prole será c/c .

De $D/d \times d/d$, metade da prole será d/d .

De $E/e \times E/e$, um quarto da prole será e/e .

Portanto, a probabilidade geral (ou frequência esperada) de se obter um genótipo $a/a ; b/b ; c/c ; d/d ; e/e$ será $\frac{1}{4} \times \frac{1}{2} \times \frac{1}{4} \times \frac{1}{2} \times \frac{1}{4} = \frac{1}{256}$. Esse cálculo de probabilidade pode ser estendido para prever as frequências fenotípicas ou frequências gaméticas. De fato, existem muitos outros usos para esse método na análise genética, e encontraremos alguns em capítulos posteriores.

Quantas proles precisamos cultivar? Para levar o exemplo precedente a uma etapa adiante, suponha ser necessário estimar quantas plantas da prole precisamos cultivar para termos uma chance razoável de obter o genótipo desejado $a/a ; b/b ; c/c ; d/d ; e/e$. Primeiro calculamos a proporção da prole que é esperada como sendo desse genótipo. Como mostrado, sabemos que precisamos examinar pelo menos 256 proles para ter uma chance média de obter uma planta individual do genótipo desejado.

A probabilidade de obter um “sucesso” (uma planta totalmente recessiva) de 256 tem que ser considerada mais cuidadosamente. Essa é a probabilidade *média* de sucesso. Infelizmente, se isolarmos e testarmos 256 da prole, muito provavelmente não teremos nenhum sucesso, mas simplesmente azar. Do ponto de vista prático, seria mais significativo perguntar: “Que tamanho de amostra precisaríamos para ter 95% de confiança de que obteríamos pelo menos um sucesso?” (**Nota:** Esse valor de 95% de confiança é padrão na ciência.) O modo mais simples de fazer esse cálculo é abordá-lo considerando a probabilidade de falha total, isto é, a probabilidade de não obter indivíduos do genótipo desejado. Em nossa amostra, para cada indivíduo isolado, a probabilidade de não ter o tipo desejado é $1 - (1/256) = 255/256$. Ampliando essa idéia para o tamanho n , vemos que a probabilidade de sucesso em uma amostra de n é $(255/256)^n$. (Essa probabilidade é uma aplicação simples da regra do produto: $255/256$ multiplicado por si mesmo n vezes.) Assim, a probabilidade de obter *pelo menos um sucesso* é a probabilidade de todos os resultados alternativos possíveis (essa probabilidade é 1) menos a probabilidade da falha total, ou $(255/256)^n$. Assim, a probabilidade de pelo menos um sucesso é $1 - (255/256)^n$. Para satisfazer o nível de confiança de 95%, devemos tornar essa expressão igual a 0,95 (o equivalente a 95%).

Portanto,

$$1 - (255/256)^n = 0,95$$

Resolvendo essa equação para n , obtemos um valor de 765, o número de prole necessário para virtualmente garantir sucesso. Note o quanto esse número é diferente na expectativa ingênua de sucesso na prole de 256. Esse tipo de cálculo é útil em muitas aplicações em genética e em outras situações nas quais um resultado bem-sucedido é necessário através de muitas tentativas.

Quantos genótipos distintos um cruzamento produz?

As regras de probabilidade podem ser facilmente usadas para prever o número de genótipos ou fenótipos na prole de linhagens parentais complexas. Por exemplo, em uma autofecundação do “tetraíbrido” $A/a ; B/b ; C/c ; D/d$, existirão três genótipos para cada par de genes; por exemplo, para o primeiro par de genes, os três genótipos serão A/a , A/A e a/a . Como existem quatro pares de genes no total, existirão $3^4 = 81$ genótipos diferentes. Em um cruzamento-teste tal como um tetraíbrido, existirão dois genótipos para cada par de genes (por exemplo, A/a e a/a) e um total de $2^4 = 16$ genótipos na prole. Como estamos supondo que todos os genes estão em cromossomos diferentes, todos esses genótipos de cruzamentos-teste ocorrerão com uma frequência igual de $\frac{1}{16}$.

Uso do teste de qui-quadrado nas proporções monoíbridas e diíbridas

Em genética, geralmente um pesquisador é confrontado com resultados que são próximos de uma proporção esperada, mas não idênticos a ela. Tais proporções podem ser de genótipos monoíbridos, diíbridos ou mais complexos, e com independência ou não. Mas quão suficientemente próximo de um resultado esperado ele é? Um teste estatístico é necessário para verificar tais proporções em relação a expectativas, e o **teste do qui-quadrado**, ou χ^2 , satisfaz esse papel.

Em que situações experimentais o teste do χ^2 é geralmente aplicável? A situação geral é aquela na qual os resultados observados são comparados com os previstos por uma hipótese. Em um exemplo genético simples, suponha que você cruza uma planta que por hipótese, com base em análise anterior, seja uma heterozigota, A/a . Para testar essa hipótese, você cruza esse heterozigoto com um testador de genótipo a/a e conta os números de fenótipos com genótipos $A/-$ e a/a na prole. Então você deve avaliar se os números que você obteve constituem a proporção esperada de 1:1. Se houver uma correlação próxima, então a hipótese é consistente com o resultado; enquanto, se houver pouca correlação, a hipótese é rejeitada. Como parte desse processo, tem que ser feito um julgamento sobre se os números observados são *suficientemente* próximos aos esperados. Correlações muito próximas e grandes diferenças em geral não apresentam problemas, mas, inevitavelmente, são áreas obscuras nas quais o ajuste não é óbvio.

O teste do χ^2 é simplesmente um modo de quantificar os vários desvios esperados pelo acaso se uma hipótese é verdadeira. Tome como exemplo a hipótese simples anterior, prevendo uma proporção de 1:1, por exemplo. Mesmo que a hipótese seja verdadeira, nem sempre esperamos uma proporção exata de 1:1. Podemos fazer um modelo dessa idéia com um barril cheio de bolas de gude vermelhas e brancas em igual número. Se retirarmos às cegas amostras de 100 bolinhas com base no acaso, podemos esperar que as amostras apresentem pequenos desvios tais como 52 vermelhas : 48 brancas muito comumente, e ter grandes desvios tais como 60 vermelhas : 40 brancas menos comumente. Até mesmo 100 bolinhas é um resultado possível, com uma probabilidade muito baixa de $(1/2)^{100}$. Entretanto, se *algum* resultado é

possível em algum nível de probabilidade mesmo que a hipótese seja verdadeira, como podemos rejeitar a hipótese? Uma convenção científica geral é que uma hipótese seria rejeitada como falsa se houver uma probabilidade de menos de 5% de observar um desvio das expectativas pelo menos tão grande quanto o observado. A hipótese ainda pode ser verdadeira, mas temos que tomar uma decisão em algum ponto, e 5% é a linha de decisão convencional. A implicação é que, embora os resultados assim tão distantes das expectativas sejam esperados em 5% das vezes mesmo quando a hipótese é verdadeira, nós, erradamente, rejeitamos a hipótese apenas em 5% dos casos e desejamos aceitar essa chance de erro. (Esses 5% são o contrário dos 95% do nível de confiança usado antes.)

Vejam os alguns dados reais. Queremos testar nossa hipótese anterior de que uma planta é heterozigota. Façamos com que *A* seja para pétalas vermelhas e *a* seja para brancas. Os cientistas testam a hipótese fazendo previsões baseadas na hipótese. Na presente situação, uma possibilidade é prever os resultados de um cruzamento-teste. Suponha que façamos um cruzamento-teste com o suposto heterozigoto. Com base na hipótese, a lei de Mendel da segregação igual prevê que devemos ter 50% *A/a* e 50% *a/a*. Suponha que, na realidade, obtemos uma prole de 120 e vimos que 55 são vermelhas e 65 são brancas. Esses números diferem das expectativas precisas, que seriam 60 vermelhas e 60 brancas. O resultado parece estar um pouco longe da proporção esperada, o que aumenta a incerteza; logo, precisamos usar o teste do χ^2 . Calculamos o χ^2 usando a seguinte fórmula:

$$\chi^2 = \sum (O - E)^2/E \text{ para todas as classes}$$

na qual *E* é o número esperado em uma classe, *O* é o número observado em uma classe e \sum significa “somatório”.

O cálculo é mais simplesmente feito usando uma tabela:

Classe	O	E	(O - E) ²	(O - E) ² /E
Vermelha	55	60	25	25/60 = 0,42
Branca	65	60	25	25/60 = 0,42
				Total = χ^2 = 0,84

Agora, temos que procurar esse valor de χ^2 no Quadro 3.1, que nos dará o valor de probabilidade que queremos. As fileiras no Quadro 3.1 mostram valores diferentes de *graus de liberdade* (*df*, *de freedom*). O número de graus de liberdade é o número de variáveis independentes nos dados. No presente contexto, o número de variáveis independentes é simplesmente o número de classes fenotípicas menos 1. Nesse caso, $df = 2 - 1 = 1$. Assim, vemos apenas a linha 1 de *df*. Vemos que nosso valor de χ^2 de 0,84 fica em algum lugar entre as colunas marcadas 0,5 e 0,1, ou seja, entre 50 e 10%. Esse valor de probabilidade é muito maior que o valor limítrofe de 5%, e assim aceitamos os resultados observados como sendo compatíveis com a hipótese.

Seguem algumas notas importantes para a aplicação desse teste:

1. O que o valor de probabilidade de fato significa? É a probabilidade de observar um desvio dos resultados esperados *pelo menos tão grande* (não *exatamente* esse desvio) com base na chance de a hipótese estar correta.
2. O fato de que nossos resultados “passaram” pelo teste do qui-quadrado porque $p > 0,05$ não significa que os

Quadro 3.1 Valores Críticos da Distribuição de χ^2

df	P									df
	0,995	0,975	0,9	0,5	0,1	0,05	0,025	0,01	0,005	
1	,000	,000	0,016	0,455	2,706	3,841	5,024	6,635	7,879	1
2	0,010	0,051	0,211	1,386	4,605	5,991	7,378	9,210	10,597	2
3	0,072	0,216	0,584	2,366	6,251	7,815	9,348	11,345	12,838	3
4	0,207	0,484	1,064	3,357	7,779	9,488	11,143	13,277	14,860	4
5	0,412	0,831	1,610	4,351	9,236	11,070	12,832	15,086	16,750	5
6	0,676	1,237	2,204	5,348	10,645	12,592	14,449	16,812	18,548	6
7	0,989	1,690	2,833	6,346	12,017	14,067	16,013	18,475	20,278	7
8	1,344	2,180	3,490	7,344	13,362	15,507	17,535	20,090	21,955	8
9	1,735	2,700	4,168	8,343	14,684	16,919	19,023	21,666	23,589	9
10	2,156	3,247	4,865	9,342	15,987	18,307	20,483	23,209	25,188	10
11	2,603	3,816	5,578	10,341	17,275	19,675	21,920	24,725	26,757	11
12	3,074	4,404	6,304	11,340	18,549	21,026	23,337	26,217	28,300	12
13	3,565	5,009	7,042	12,340	19,812	22,362	24,736	27,688	29,819	13
14	4,075	5,629	7,790	13,339	21,064	23,685	26,119	29,141	31,319	14
15	4,601	6,262	8,547	14,339	22,307	24,996	27,488	30,578	32,801	15

resultados são compatíveis com a hipótese. Entretanto, se tivéssemos obtido um valor $p < 0,05$, seríamos forçados a rejeitar a hipótese. A ciência é sobre hipóteses falsificáveis, não a “verdade”.

3. Devemos ter cuidado com as palavras sobre as hipóteses, pois suposições tácitas em geral são ocultas nelas. A hipótese atual é um caso assim. Se formos expressá-la cuidadosamente, teríamos que dizer que o “indivíduo em teste é um heterozigoto A/a , que esses alelos mostram segregação igual na meiose e que as proles A/a e a/a são igualmente viáveis”. Investigaremos os efeitos dos alelos sobre a viabilidade no Capítulo 6, mas, no momento, devemos ter em mente a possível complicação porque as diferenças na sobrevivência afetariam os tamanhos das várias classes. O problema é que, se rejeitamos uma hipótese que escondeu componentes, não sabemos quais dos componentes estamos rejeitando. Por exemplo, no presente caso, se somos forçados a rejeitar a hipótese como resultado do teste de χ^2 , não saberemos se estamos rejeitando a segregação igual ou a viabilidade igual ou ambas.
4. O resultado do teste do χ^2 depende muito do tamanho das amostras (números nas classes). Assim, o teste deve usar *números reais*, não proporções ou porcentagens. Adicionalmente, quanto maiores as amostras, mais confiável é o teste.

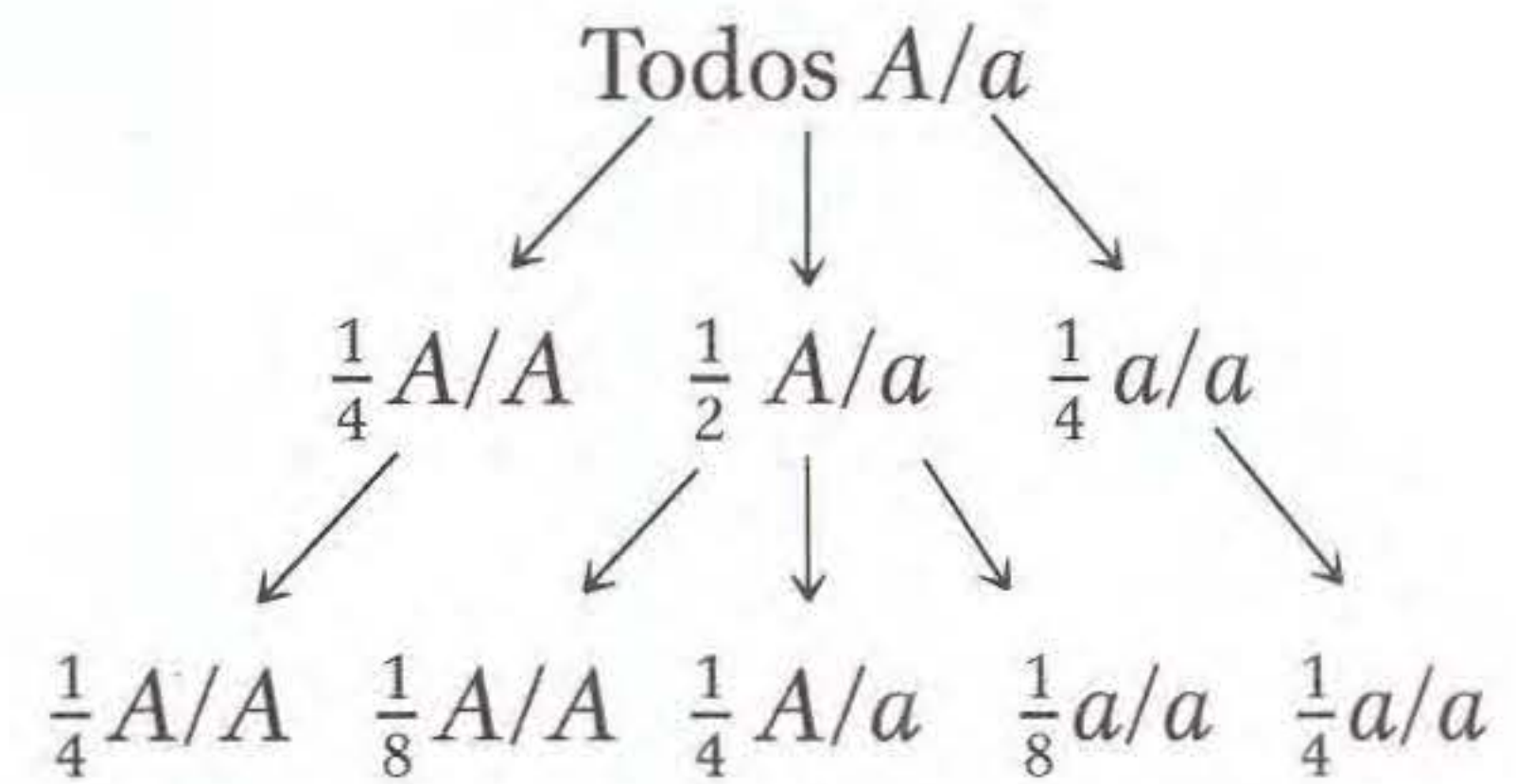
Qualquer das proporções mendelianas familiares neste capítulo ou no Capítulo 2 pode ser testada usando o teste do χ^2 — por exemplo, 3:1 (1 df), 1:2:1 (2 df), 9:3:3:1 (3 df) e 1:1:1:1 (3 df). Voltaremos para mais explicações do teste de χ^2 no Capítulo 4.

Síntese de linhagens puras

As linhagens puras estão entre as ferramentas essenciais da genética. Por um lado, apenas essas linhagens totalmente homozigotas expressarão alelos recessivos, mas a principal necessidade de linhagens puras é a manutenção de estoques para pesquisa. Os membros de uma linhagem pura podem ser entrecruzados com o tempo e, assim, agir como uma fonte constante de genótipo para uso em experimentos. Assim, para a maioria dos organismos-modelo, existem centros de estoques internacionais que são repositórios de linhagens puras. Centros similares de estoques fornecem linhagens de plantas e animais para uso na agricultura.

As linhagens puras de plantas ou animais são feitas por gerações sucessivas de autofecundação. (Nos animais, a autofecundação é feita cruzando animais de genótipos idênticos.) A autofecundação de uma planta monoíbrida mostra o princípio em funcionamento. Suponha que começamos com uma população de indivíduos que são todos A/a e deixamos que ocorra autofecundação. Podemos aplicar a Primeira Lei de Mendel para prever que, na geração seguinte, teremos $\frac{1}{4} A/A$, $\frac{1}{2} A/a$ e $\frac{1}{4} a/a$. Note que a heterozigose

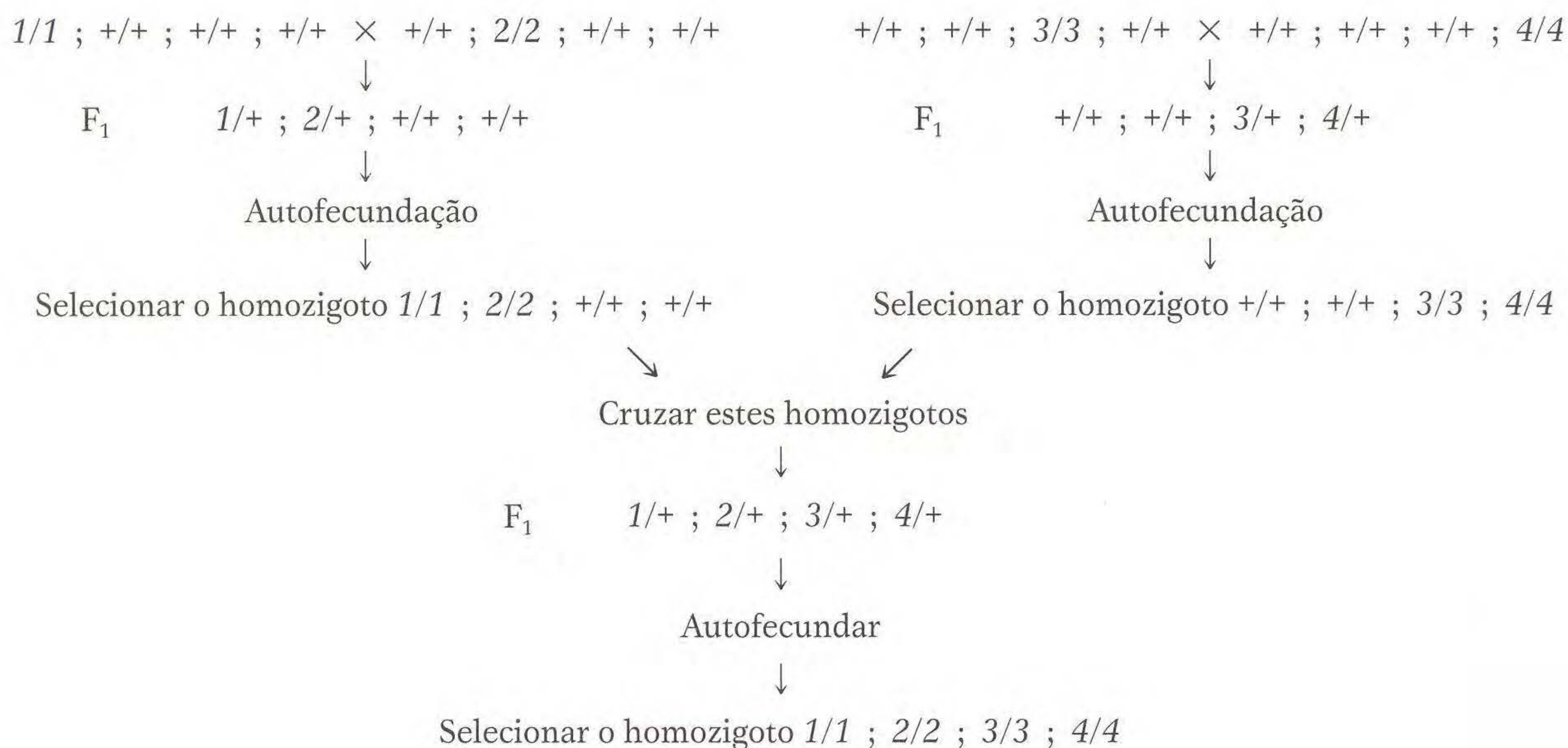
(a proporção de heterozigotos) foi reduzida à metade, de 1 para $\frac{1}{2}$. Se repetirmos esse processo de autofecundação para outra geração, todos os descendentes dos homozigotos serão homozigotos, mas, novamente, os heterozigotos terão sua proporção de um quarto. O processo é mostrado do seguinte modo:



Após, digamos, oito gerações de autofecundação, a proporção de heterozigotos é reduzida a $(1/2)^8$, que é $1/256$, ou cerca de 0,4%. Vejamos esse processo de um modo ligeiramente diferente: vamos supor que começemos tal programa com um genótipo que é heterozigoto em 256 pares de genes. Se também for suposto haver distribuição independente, então, após a autofecundação por oito gerações, terminaremos com uma gama de genótipos, cada um tendo, em média, apenas um gene heterozigoto (isto é, $1/256$). Em outras palavras, estaremos no caminho para criar um número de linhagens puras.

Vamos aplicar esse princípio para a seleção de linhagens agrícolas, o tópico com o qual começamos o capítulo. Podemos usar como nosso exemplo a seleção de trigo Marquis por Charles Saunders na primeira parte do século XX. A meta de Saunders era desenvolver uma linhagem produtiva de trigo que tivesse um período curto de crescimento e ocupasse grandes áreas de terra nos países do norte, tais como Canadá e Rússia, para cultivar trigo, outro alimento principal do mundo. Ele cruzou uma linhagem tendo excelente qualidade de grãos chamada Red Fife com uma linhagem chamada Hard Red Calcutta, que, mesmo tendo produtividade e qualidade pobres, amadurecia 20 dias mais cedo que a Red Fife. A F_1 produzida pelo cruzamento foi supostamente multiplicar heterozigotos para os genes que controlam as qualidades do trigo. A partir dessa F_1 , Saunders fez autofecundações e seleções que levaram a uma linhagem pura que tinha a combinação de propriedades favoráveis necessárias para grãos de boa qualidade e maturação mais cedo. Essa linhagem foi chamada de Marquis, tendo sido rapidamente adotada em muitas partes do mundo.

Um enfoque similar pode ser aplicado a linhagens de arroz com a qual começamos o capítulo. Todas as mutações monogênicas são cruzadas em pares, e então suas plantas F_1 são autofecundadas ou entrecruzadas com outras plantas de F_1 . Como demonstração, consideremos apenas quatro mutações, de 1 a 4. Um programa de cruzamentos pode ser o seguinte, no qual os alelos mutantes e suas contrapartes tipo selvagem são sempre citadas na mesma ordem:



Esse tipo de cruzamento foi aplicado a muitas outras espécies de cultivo. As diversas linhagens puras e coloridas de tomates usadas no comércio são mostradas na Figura 3.5.

Note que, em geral quando um heterozigoto múltiplo é autofecundado, é produzida uma gama de homozigotos diferentes. Por exemplo, de $A/a ; B/b ; C/c$ existem dois homozigotos para cada par de genes (isto é, para o primeiro gene, os homozigotos são A/A e a/a), e assim existem $2^3 = 8$ homozigotos diferentes possíveis: $A/A ; b/b ; C/c$, e $a/a ; B/B ; c/c$, e assim por diante. Cada homozigoto distinto pode começar uma nova linhagem pura.

As autofecundações repetidas levam a uma proporção aumentada de homozigotos, um processo que pode ser usado para criar linhagens puras para pesquisa ou outras aplicações.

Vigor híbrido

Consideramos a síntese de linhagens puras superiores para pesquisa e para agricultura. As linhagens puras são convenientes, pois a propagação do genótipo de ano a ano é bastante fácil. Entretanto, uma grande proporção de sementes comerciais que os fazendeiros (e jardineiros) usam é chamada de *semente híbrida*. Curiosamente, em muitos casos nos quais duas linhagens diferentes de plantas (e animais) são unidas em um híbrido de F_1 (suposto heterozigoto), o híbrido apresenta tamanho e vigor maiores do que as duas linhagens contribuintes (Figura 3.6). Essa superioridade geral de múltiplos heterozigotos é chamada de **vigor híbrido**. Os motivos moleculares para o vigor híbrido são praticamente desconhecidos e ainda intensamente debatidos, mas o fenômeno é inegável e tem feito grandes contribuições para a agricultura. Um aspecto negativo de usar híbridos é que, a cada estação, as duas linhagens parentais devem ser cultivadas separadamente e, então, entrecruzadas para fazer a semente híbrida para venda. Esse processo é muito mais inconveniente do que manter linhagens puras, que necessitam apenas deixar que as plantas se autofecundem; conseqüentemente, a semente híbrida é mais cara do que a de linhagens puras.

Na perspectiva dos usuários, há um outro aspecto negativo de usar híbridos. Após uma planta híbrida ter crescido e produzido sua colheita para venda, não é realista manter algumas das sementes que ela produz e esperar que elas sejam igualmente vigorosas no ano seguinte. O motivo é que, quando o híbrido sofre meiose, a distribuição independente dos vários pares de genes mistos formará muitas combinações alélicas diferentes, e poucas dessas combinações serão do híbrido original. Por exemplo, o tetraíbrido já descrito, quando autofecundado, produz 81 genótipos diferentes, dos quais apenas uma minoria será tetraíbrida. Supondo-se a distribuição independente, então, para cada par de gene, a autofecundação produziria metade de heterozigotos ($A/a \rightarrow \frac{1}{4} A/A, \frac{1}{2} A/a$ e $\frac{1}{4} a/a$). Como existem quatro pares de genes nes-



FIG. 3.5 O cruzamento de tomates resultou em uma grande gama de linhagens de genótipos e fenótipos diferentes. [David Cavagnaro/Visuals Unlimited.]

Vigor híbrido em milho



(b)

(a)

FIG. 3.6 Vários heterozigotos híbridos flanqueados por duas linhagens puras cruzadas para fazê-los. (a) As plantas. (b) Espigas das mesmas plantas. [(a) Cortesia de Ruth A. Swanson-Wagner, Schnable Lab, Iowa State University; (b) cortesia de Jun Cao, Schnable Lab, Iowa State University.]

se tetraíbrido, a proporção da prole que será igual ao híbrido original $A/a ; B/b ; C/c ; D/d$ será $(1/2)^4 = 1/16$.

Alguns híbridos entre linhagens geneticamente diferentes apresentam vigor híbrido. Entretanto, a distribuição de genes quando o híbrido sofre meiose desfaz a combinação alélica favorável, e assim alguns membros da geração seguinte a têm.

► 3.3 A Base Cromossômica da Distribuição Independente

Como a segregação igual, a distribuição independente de pares de genes em cromossomos diferentes é explicada pelo comportamento dos cromossomos durante a meiose. Considere um cromossomo, que poderemos chamar de número 1; seus dois homólogos podem ser chamados de $1'$ e $1''$. Se os cromossomos se alinham no equador, então $1'$ pode ir para o “norte” e $1''$ para o “sul”, ou vice-versa. Similarmente, para o cromossomo 2 com homólogos $2'$ e $2''$, $2'$ pode ir para o norte e $2''$ para o sul, ou vice-versa. Assim, o cromossomo $1'$ pode terminar embalado com o cromossomo $2'$ ou $2''$, dependendo de quais cromossomos foram levados na mesma direção.

A distribuição independente não é fácil de demonstrar observando cromossomos segregantes ao microscópio, pois os homólogos tais como $1'$ e $1''$ geralmente não parecem diferentes, embora eles possam levar uma pequena sequência de variação. Entretanto, a distribuição independente pode ser observada em alguns casos especializados. Um caso foi instrumental no desenvolvimento histórico da teoria cromossômica.

Em 1913, Elinor Carothers encontrou uma situação cromossômica incomum em certa espécie de gafanhoto, uma situação que permite um teste direto se pares diferentes de cromossomos de fato segregam independentemente. Estudando a meiose nos testículos de gafanhotos, ela descobriu um gafanhoto no qual um “par de cromossomos” tinha membros não-idênticos. Tal par é chamado de par *heteromórfico*; supostamente, os cromossomos apresentam apenas homologia parcial. Além disso, o mesmo gafanhoto tinha outro cromossomo (não relacionado ao par heteromórfico) que não tinha parceiro de pareamento. Carothers foi capaz de usar esses cromossomos incomuns como marcadores citológicos visíveis do comportamento de cromossomos durante a meiose. Observando muitas meioses, ela pôde contar o número de vezes que determinado membro do par heteromórfico migrou para o mesmo pólo que o cromossomo sem parceiro de pareamento (Figura 3.7). Ela observou os dois padrões de comportamento cromossômico com frequência igual. Embora esses cromossomos incomuns não sejam típicos, os resultados sugerem fortemente que, em geral, cromossomos não-homólogos distribuem-se independentemente.

Distribuição independente em organismos diplóides

A base cromossômica da lei da distribuição independente é formalmente diagramada na Figura 3.8, que ilustra como o comportamento separado de dois pares de cromossomos diferentes dá origem a proporções mendelianas de 1:1:1:1 dos tipos gaméticos esperados pela distribuição independente. A célula hipotética tem quatro cromossomos: um par de cromossomos longos homólogos e um par de cromossomos

Cromossomos diferentes segregam-se independentemente

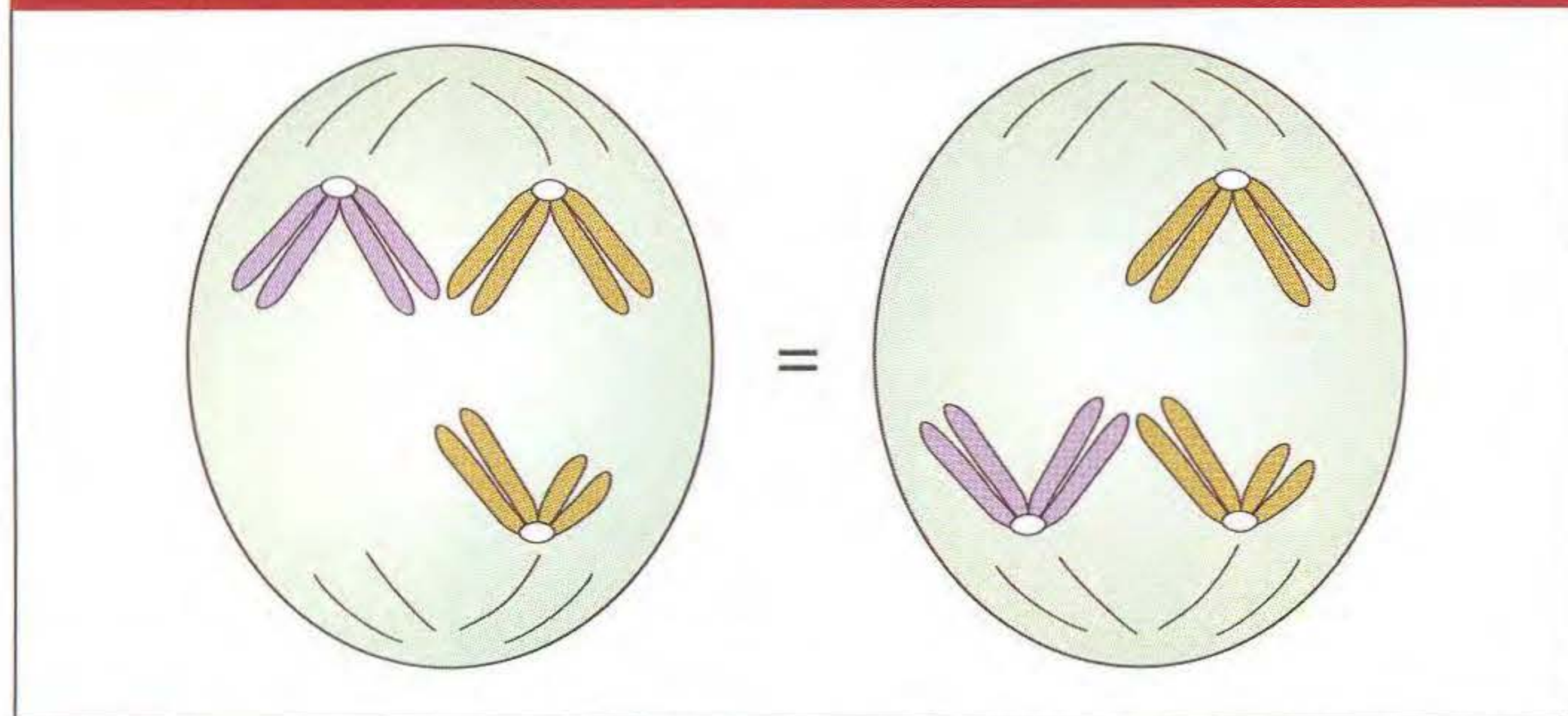


FIG. 3.7 Carothers observou esses dois padrões igualmente frequentes pelos quais um par heteromórfico e um cromossomo não-pareado movem-se para gametas na meiose.

homólogos pequenos. O genótipo dos meiócitos é $A/a ; B/b$, e os dois pares alélicos, A/a e B/b , são mostrados em dois pares de cromossomos diferentes. As partes 4 e 4' na Figura 3.8 mostram uma etapa-chave na distribuição independente: existem dois padrões de segregação alélica igualmente frequentes, um mostrado em 4 e outro em 4'. Em um caso, os alelos A/A e B/B são levados juntos para uma célula, e a/a e b/b juntos para outra célula. No outro caso, os alelos A/A e b/b são unidos na mesma célula. Os dois padrões resultam de duas ligações ao fuso igualmente frequentes dos centrômeros na anáfase I. A meiose, então, produz quatro células dos genótipos indicados de cada um desses padrões de segregação. Como os padrões de segregação 4 e 4' são igualmente comuns, as células de produtos meióticos com os genótipos $A ; B ; a ; b$, $A ; b$ e $a ; B$ são produzidas em iguais frequências. Em outras palavras, a frequência de cada um dos quatro genótipos é $\frac{1}{4}$. A distribuição gamética é a postulada por Mendel para um diíbrido, e nós a inserimos na margem do quadrado de Punnett. A fusão aleatória desses gametas resulta na proporção fenotípica 9:3:3:1 da F_2 .

Distribuição independente em organismos haplóides

Nos fungos ascomicetos, podemos inspecionar os produtos de um único meiócito para mostrar diretamente a distribuição independente. Vamos usar o fungo filamentoso *Neurospora crassa* para ilustrar esse ponto (veja o boxe Organismo-modelo adiante). Como já vimos nos exemplos anteriores de fungos, é feito um cruzamento em *Neurospora* misturando duas linhagens parentais haplóides de tipo reprodutivo oposto. Em um modo similar ao de leveduras, o tipo reprodutivo é determinado por dois "alelos" de um gene, nessa espécie, chamados de MAT-A e MAT-a. O modo pelo qual o cruzamento é feito é mostrado na Figura 3.9.

Os produtos da meiose em fungos são esporos sexuais. Lembre que os *ascomicetos* (que incluem *Neurospora* e *Saccharomyces*) são únicos pelo fato de que, para qualquer determinado meiócito, os esporos são mantidos juntos em um saco membranoso chamado *asco*. Assim, para esses organismos, os produtos de uma única meiose podem ser recuperados e

testados. No mofo do pão *Neurospora*, os fusos nucleares da meiose I e II não se superpõem dentro do asco em forma de cigarro, e, portanto, os quatro produtos de um único meiócito ficam em uma linha reta (Figura 3.10a). Além disso, por algum motivo não compreendido, há uma *mitose pós-meiótica*, que também não apresenta superposição do fuso. Logo, a meiose e a mitose extra resultam em um asco linear contendo oito *ascosporos*. Em um meiócito heterozigoto A/a , se não houver crossings entre o gene e seu centrômero (veja Capítulo 4), então existirão dois blocos adjacentes de ascosporos, quatro de A e quatro de a (Figura 3.10b).

Agora podemos examinar um diíbrido. Façamos um cruzamento entre dois mutantes distintos tendo mutações em genes diferentes em cromossomos diferentes. Supondo que os loci dos genes mutantes estão ambos muito próximos de seus centrômeros respectivos, evitamos complicações devidas ao crossing over entre os loci e os centrômeros. O primeiro mutante é albino (a), contrastando com o tipo selvagem rosa (a^+). O segundo mutante é um biscoito (b), que tem uma colônia muito compacta em forma de biscoito em contraste com a colônia achatada do tipo selvagem (b^+). Vamos supor que os dois mutantes são do tipo reprodutivo oposto. Assim, o cruzamento é

$$a ; b^+ \times a^+ ; b$$

Devido à ligação aleatória ao fuso, os seguintes dois tipos de óctades serão igualmente frequentes:

$a^+ ; b$	$a ; b$
$a^+ ; b$	$a ; b$
$a^+ ; b$	$a ; b$
$a^+ ; b$	$a ; b$
$a ; b^+$	$a^+ ; b^+$
$a ; b^+$	$a^+ ; b^+$
$a ; b^+$	$a^+ ; b^+$
$a ; b^+$	$a^+ ; b^+$
50%	50%

A frequência igual desses dois tipos é uma demonstração convincente da distribuição independente que ocorre em meiócitos individuais.

Distribuição independente de genes autossômicos e ligados ao X

O princípio da distribuição independente também é útil na análise de genótipos que são heterozigotos para genes autossômicos e ligados ao X. Os autossomos e os cromossomos sexuais são movidos independentemente por fibras do fuso ligadas aleatoriamente a seus centrômeros, como com dois pares diferentes de autossomos. Algumas proporções diíbridas interessantes são produzidas. Vejamos um exemplo de *Drosophila*. O cruzamento é feito entre uma fêmea com asas vestigiais (autossômico recessivo, vg) e um macho com olhos brancos (recessivo ligado ao X, w). Simbolicamente, o cruzamento é

$$vg/vg ; +/+ \text{ ♀ } \times +/+ ; w/Y \text{ ♂ }$$

Distribuição independente de cromossomos na meiose explica a proporção de Mendel

Recombinação meiótica entre genes não-ligados por segregação independente

W.W. ANIMAÇÃO COMPUTORIZADA

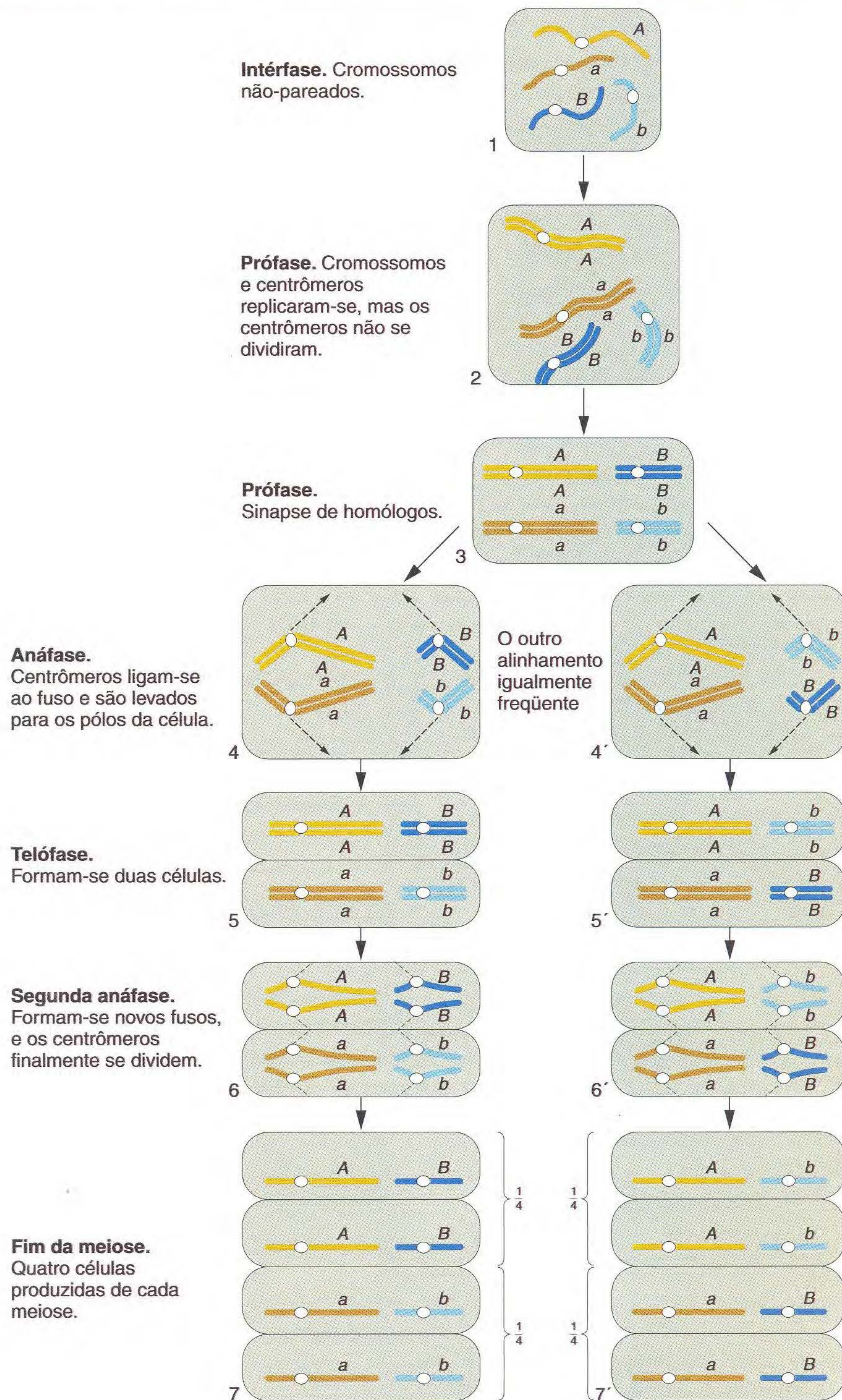


FIG. 3.8 Meiose em uma célula diplóide com genótipo $A/a ; B/b$. O diagrama mostra como a segregação e a distribuição de pares cromossômicos diferentes dão origem à proporção gamética mendeliana 1:1:1:1.

Estágios de um cruzamento de *Neurospora*

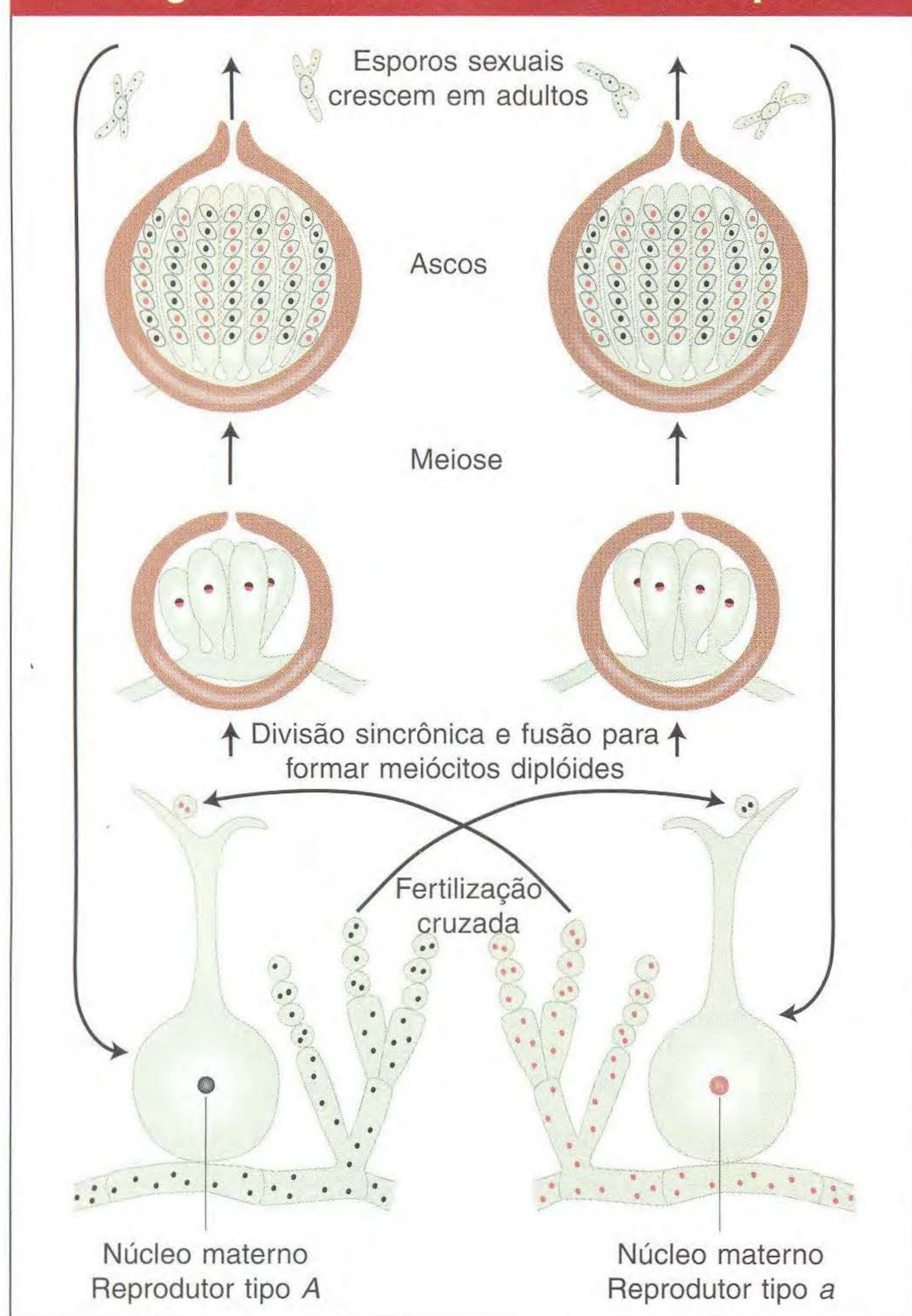


FIG. 3.9 O ciclo de vida de *Neurospora crassa*, o mofo laranja do pão. A autofertilização não é possível nessa espécie: existem dois tipos reprodutivos, determinados pelos alelos A e a de um gene, e qualquer um deles pode agir como uma "fêmea". Um esporo assexual do tipo reprodutivo oposto funde-se com um filamento receptivo, e um núcleo do esporo assexual viaja pelo filamento até parear com um núcleo feminino nas células. O par A e a então sofre mitoses sincrônicas, finalmente fundindo-se para formar meiospóides.

A F_1 será:

Fêmeas com genótipo $+/vg ; +/w$
 Machos com genótipo $+/vg ; +/Y$

Essas moscas F_1 devem ser entrecruzadas para obter uma F_2 . Como o cruzamento é monóibrido para o gene autossômico *vestigial*, ambos os sexos da F_2 mostrarão

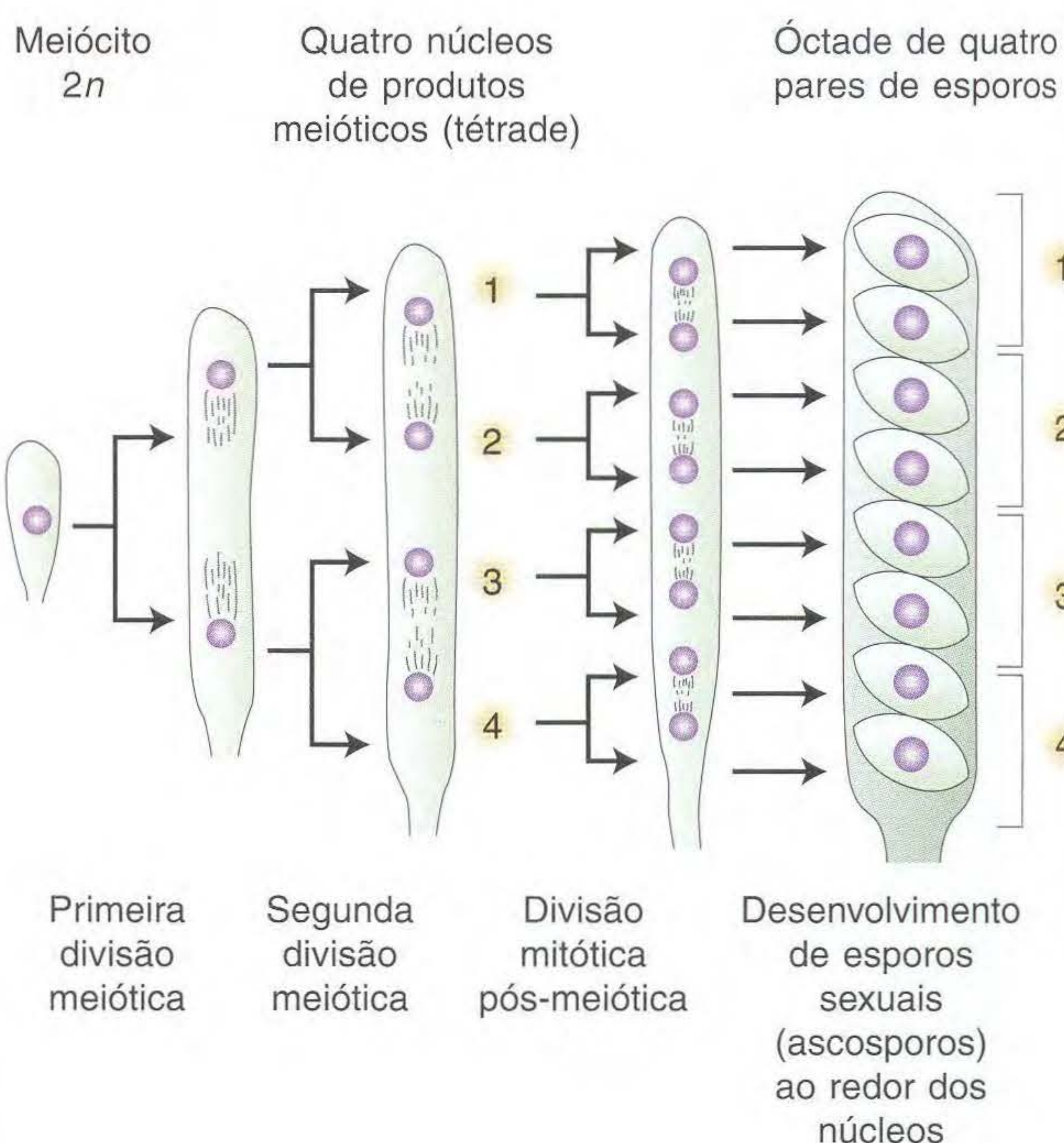
Fêmeas e machos $\frac{3}{4} +/-$ (tipo selvagem)
 $\frac{1}{4} vg/vg$ (vestigial)

Para o gene ligado ao X *white*, as proporções serão as seguintes:

Fêmeas $\frac{1}{2} ++$ e $\frac{1}{2} +/w$ (todos tipo selvagem)
 Machos $\frac{1}{2} +/Y$ (tipo selvagem) e $\frac{1}{2} w/Y$ (white)

A meiose linear de *Neurospora*

(a) Divisões nucleares



(b) Segregação de alelos

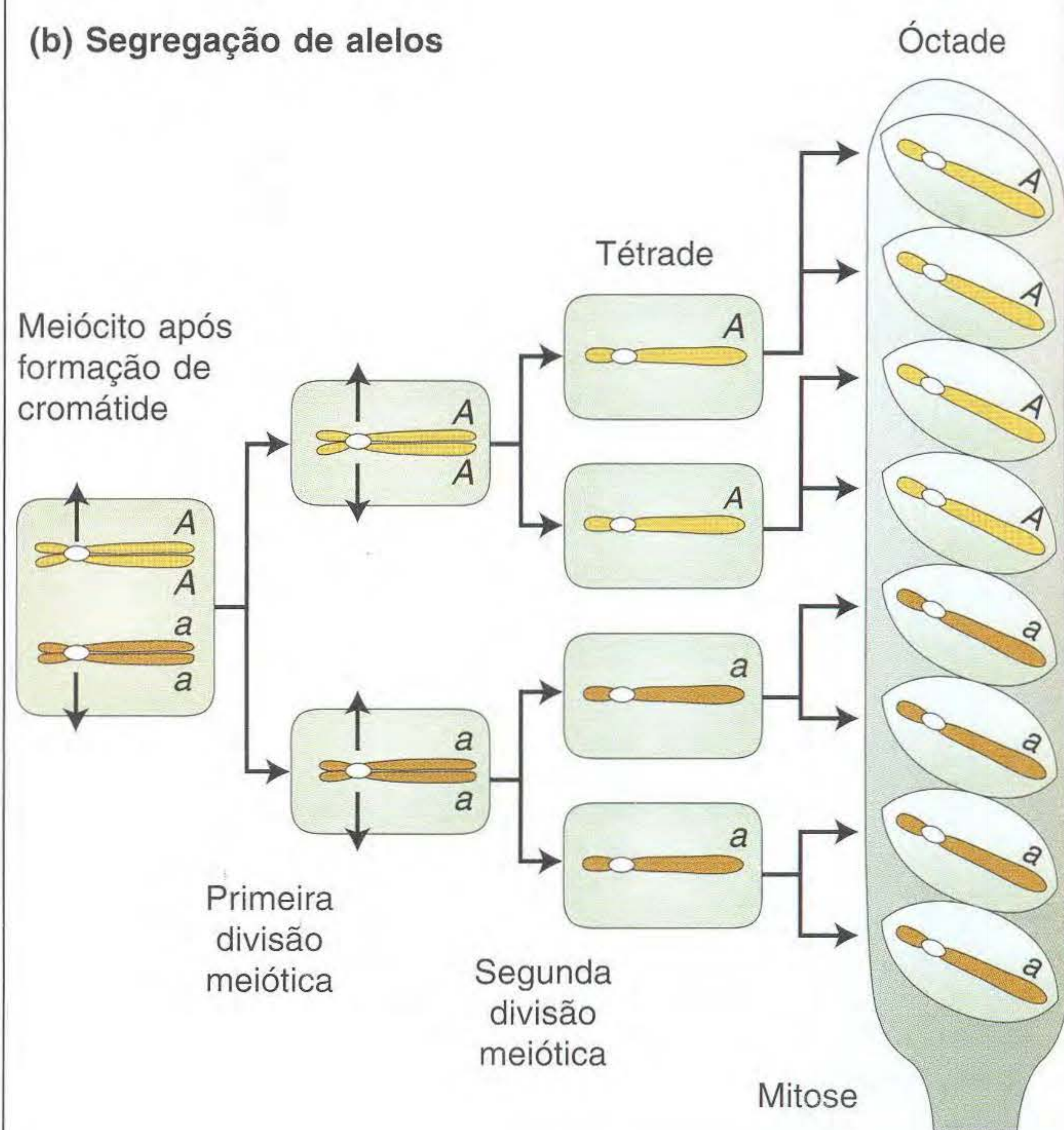


FIG. 3.10 *Neurospora* é um sistema-modelo ideal para estudar a segregação alélica na meiose. (a) Os quatro produtos da meiose (tétrade) sofrem mitose para produzir uma óctade. Os produtos estão contidos dentro de um asco. (b) Um meiospóide A/a sofre meiose seguida de mitose, resultando em números iguais de produtos A e a e demonstrando o princípio de segregação igual.

Se os genes autossômicos e ligados ao X são combinados, as proporções fenotípicas de F_2 serão

Fêmeas	$\frac{3}{4}$ totalmente tipo selvagem $\frac{1}{4}$ vestigial
Machos	$\frac{3}{8}$ totalmente tipo selvagem ($\frac{3}{4} \times \frac{1}{2}$) $\frac{3}{8}$ white ($\frac{3}{4} \times \frac{1}{2}$) $\frac{1}{8}$ vestigial ($\frac{1}{4} \times \frac{1}{2}$) $\frac{1}{8}$ vestigial, white ($\frac{1}{4} \times \frac{1}{2}$)

Assim, vemos uma proporção de prole que revela elementos claros de ambas as heranças, autossômica e ligada ao X.

Recombinação

A distribuição independente dos genes na meiose é um dos principais modos pelos quais um organismo produz novas combinações de alelos. A produção de novas combinações de alelos é formalmente chamada de **recombinação**.

Há uma concordância geral de que o motivo pelo qual esses organismos produzem novas combinações de alelos é dar variação como matéria-prima para a seleção natural. A recombinação é um princípio crucial em genética, em parte devido à sua relevância para a evolução, mas também devido a seu uso na análise genética. Ela é particularmente útil para analisar os padrões de herança de genótipos multigênicos.



Organismo-modelo *Neurospora*

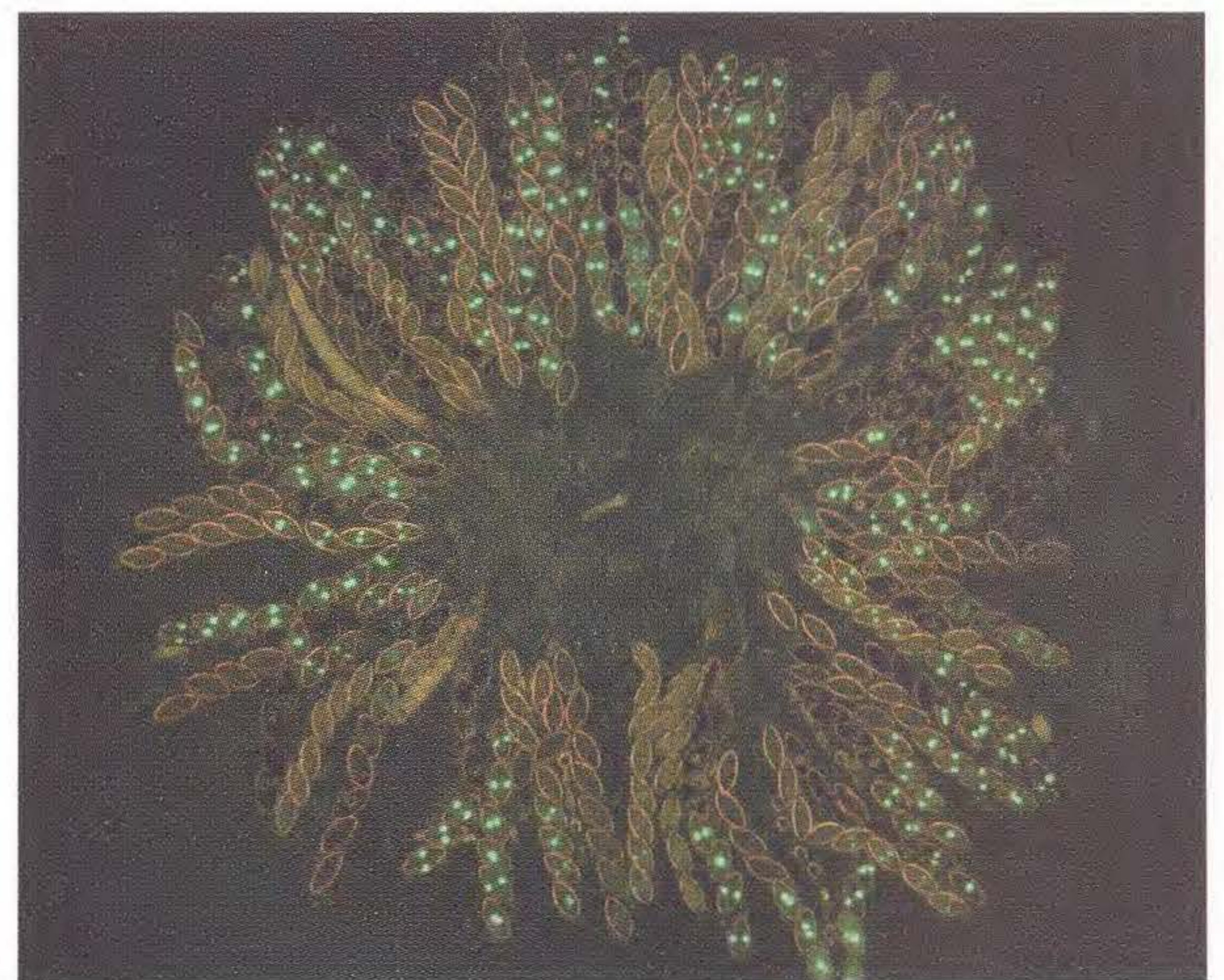
A *Neurospora crassa* foi um dos primeiros micróbios eucarióticos a serem adotados pelos geneticistas como um organismo-modelo. É um fungo haplóide ($n = 7$) que cresce em vegetação morta em muitas partes do mundo. Quando um esporo assexual (haplóide) germina, ele produz uma estrutura tubular que se estende rapidamente por crescimento da ponta e cria múltiplos ramos laterais. O resultado é uma massa de filamentos ramificados (chamados *hifas*), que constituem uma colônia. As hifas não têm paredes cruzadas, e assim a colônia é essencialmente uma célula contendo muitos núcleos haplóides. De uma colônia brotam milhões de esporos assexuais, que podem dispersar-se e repetir o ciclo assexual.

As colônias assexuais são mantidas facilmente e com baixo custo no laboratório em um meio definido de sais inorgânicos mais uma fonte de energia tal como açúcar. (Um gel inerte tal como ágar é adicionado para dar uma superfície firme.) O fato de que a *Neurospora* pode produzir quimicamente todas as suas moléculas essenciais a partir de um meio tão simples levou os geneticistas bioquímicos (começando com George Beadle e Edward Tatum; veja Capítulo 6) a preferi-la para estudos de vias sintéticas. Os geneticistas desenvolveram as etapas nessas vias introduzindo mutações e observando seus efeitos. O estado haplóide da *Neurospora* é ideal para tais análises mutacionais porque os alelos mutantes são sempre expressos diretamente no fenótipo.

A *Neurospora* tem dois tipos reprodutivos, MAT-A e MAT-a, que podem ser vistos como simples “sexos”. Quando as colônias de tipos reprodutivos diferentes entram em contato, suas paredes celulares e núcleos se fundem, resultando em muitos núcleos diplóides transitórios, cada um dos quais sofre meiose. Os quatro produtos haplóides da meiose ficam juntos em um saco chamado *asco*. Cada um desses produtos da meiose sofre outra divisão mitótica, resultando em oito ascósporos dentro de cada asco. Os ascósporos germinam e produzem colônias exatamente iguais às produzidas por esporos assexuais. Assim, tais fungos *ascomycetos* são ideais para o estudo da segregação e recombinação de genes em meioses individuais.



(a)



(b)

O fungo *Neurospora crassa*. (a) Colônias laranja de *Neurospora crassa* crescem em cana-de-açúcar. Na natureza, colônias de *Neurospora* são mais frequentemente encontradas após um incêndio, que ativa ascósporos dormentes. (Os canaviais são queimados para remover a folhagem antes da colheita da cana.) (b) Desenvolvimento de óctades de *Neurospora* de um cruzamento do tipo selvagem com uma linhagem levando um alelo modificado de uma proteína fluorescente verde de água viva fusionado com histona. As óctades apresentam a esperada segregação mendeliana 4:4 da fluorescência. Em alguns esporos, o núcleo dividiu-se mitoticamente para formar dois, e cada esporo conterá vários núcleos. [(a) Cortesia de David Jacobson; (b) cortesia de Namboori B. Raju.]

Nesta seção, definimos a recombinação de tal modo que a reconheceríamos em resultados experimentais, bem como o modo pelo qual a recombinação é analisada e interpretada.

A recombinação é observada em uma variedade de situações biológicas, mas, no presente, nós a definimos em relação à meiose.

Recombinação meiótica é definida como qualquer processo meiótico que gera um produto haplóide com novas combinações de alelos levados pelos genótipos haplóides que se unem para formar o meiócito.

Essa definição aparentemente extensa é de fato muito simples; ela destaca o ponto importante de que detectamos a recombinação comparando as *entradas* na meiose com os *resultados* (Figura 3.11). As entradas são os dois genótipos haplóides que se combinam para formar o meiócito, a célula diplóide que sofre meiose. Para os humanos, as entradas (*inputs*) são o ovócito e o espermatozóide parental. Eles se unem para formar um zigoto diplóide, que se divide para produzir todas as células do corpo, incluindo os meiócitos que são postos de lado dentro das gônadas. Os genótipos resultantes são os produtos haplóides da meiose. Nos humanos, esses produtos haplóides são os próprios ovócitos de uma pessoa, ou espermatozoides. Qualquer produto meiótico que tenha uma nova combinação de alelos dada pelos dois genótipos de entrada é por definição um **recombinante**.

A meiose gera recombinantes, que são produtos meióticos haplóides com novas combinações de alelos levados pelos genótipos haplóides que se unem para formar o meiócito.

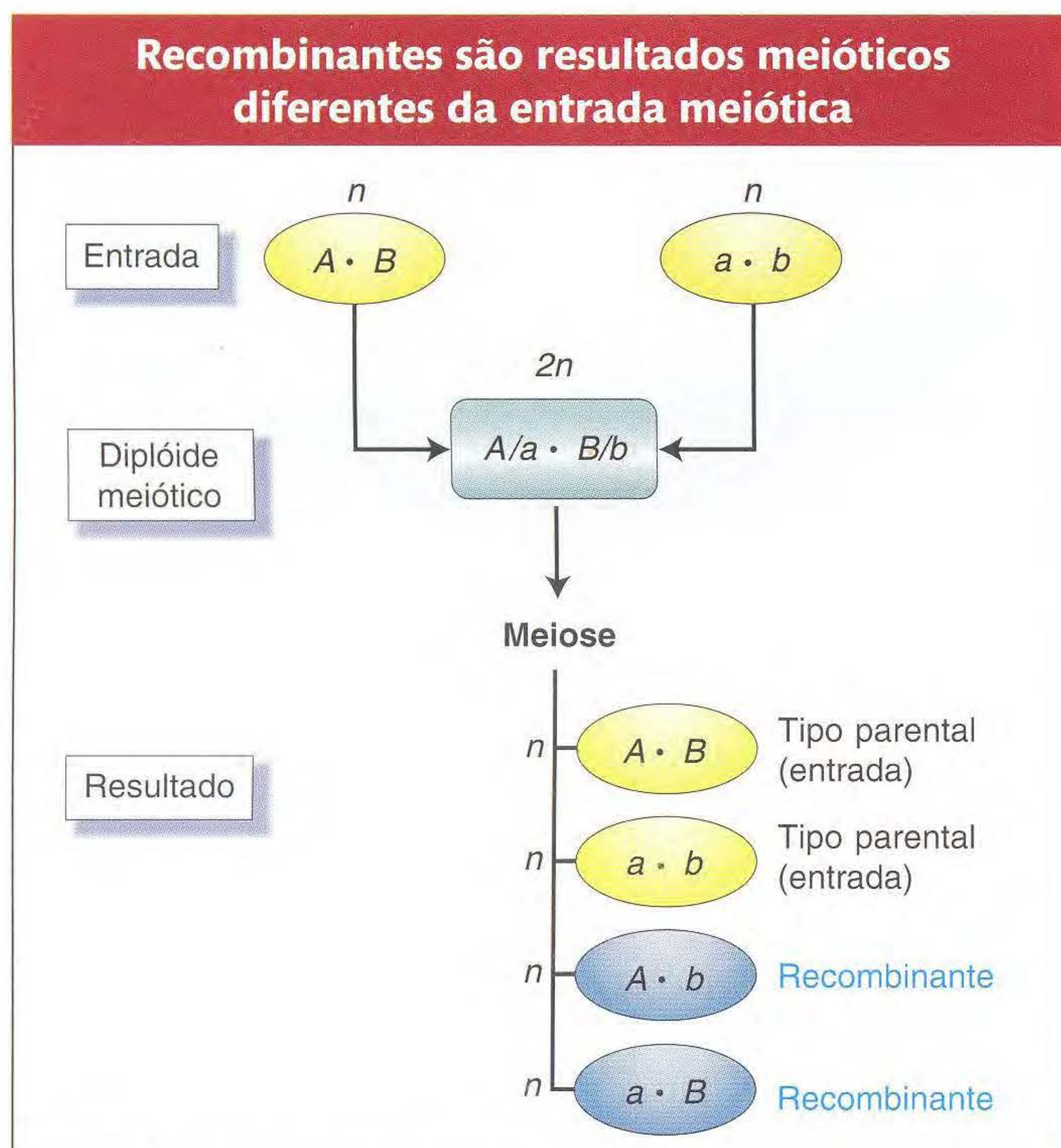


FIG. 3.11 Recombinantes são os produtos da meiose com combinações alélicas diferentes das de células haplóides que formaram o diplóide meiótico.

Primeiro, vejamos como os recombinantes são detectados experimentalmente. A detecção de recombinantes em organismos com ciclos de vida haplóides, tais como fungos e algas, é direta. Os tipos que entram nos e saem dos ciclos de vida haplóides são os genótipos de indivíduos ao invés de gametas, e podem portanto ser deduzidos diretamente dos fenótipos. A Figura 3.11 pode ser vista como um resumo da simples detecção de organismos recombinantes com ciclos de vida haplóides. A detecção de recombinantes em organismos com ciclos de vida haplóides é ardilosa. Os tipos que entram nos e saem dos ciclos diplóides são gametas. Assim, devemos conhecer os genótipos de gametas de entrada e saída para detectar recombinantes em um organismo com um ciclo diplóide. Não podemos detectar os genótipos de gametas que entram e saem diretamente, mas podemos deduzir esses genótipos usando as técnicas apropriadas:

- Para conhecer os gametas que entram (*input*), usamos genitores diplóides de reprodução pura porque eles podem produzir apenas um tipo gamético.
- Para detectar os gametas recombinantes produzidos, fazemos um cruzamento-teste com o indivíduo diplóide e observamos sua prole (Figura 3.12).

Uma prole de cruzamento-teste que surge de um produto recombinante de meiose também é chamada de um *recombinante*. Note, novamente, que o cruzamento-teste nos permite concentrar em uma meiose e evitar a ambigüidade. A partir de uma *autofecundação* da F_1 na Figura 3.10, por exemplo, a prole recombinante $A/A \cdot B/b$ não pode ser distinta de $A/A \cdot B/B$ sem outros cruzamentos.

A parte central da análise de recombinação é a frequência de recombinação. Um motivo para focar a frequência de recombinantes é que seu valor numérico é um teste convencional para ver se dois genes estão em cromossomos diferentes. Os recombinantes são produzidos por dois processos celulares diferentes: a distribuição independente de genes em cromossomos diferentes (este capítulo) e o *crossing over* entre genes no mesmo cromossomo (veja Capítulo 4). A *proporção* de recombinantes é a idéia principal aqui porque o valor diagnóstico pode nos dizer se os genes estão em cromossomos diferentes. Lidaremos com a distribuição independente aqui.

Para genes em cromossomos separados, os recombinantes são produzidos por distribuição independente, como mostrado na Figura 3.13. Novamente, vemos a proporção 1:1:1:1 que já vimos antes, mas agora a prole do cruzamento-teste é classificada ou como recombinante, ou como similar aos tipos de entrada (parentais) P. Dito desse modo, a proporção de recombinantes é claramente $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, ou 50% da prole total. Assim, vemos que a distribuição independente na meiose produz frequência de recombinantes de 50%. Se observamos uma frequência de recombinantes de 50% em um cruzamento-teste, podemos deduzir que os dois genes em estudo distribuem-se independentemente. A interpretação mais simples e mais provável da distribuição independente é que os dois genes estão em pares de cromossomos separados. (Entretanto, devemos notar que os genes que estão bem

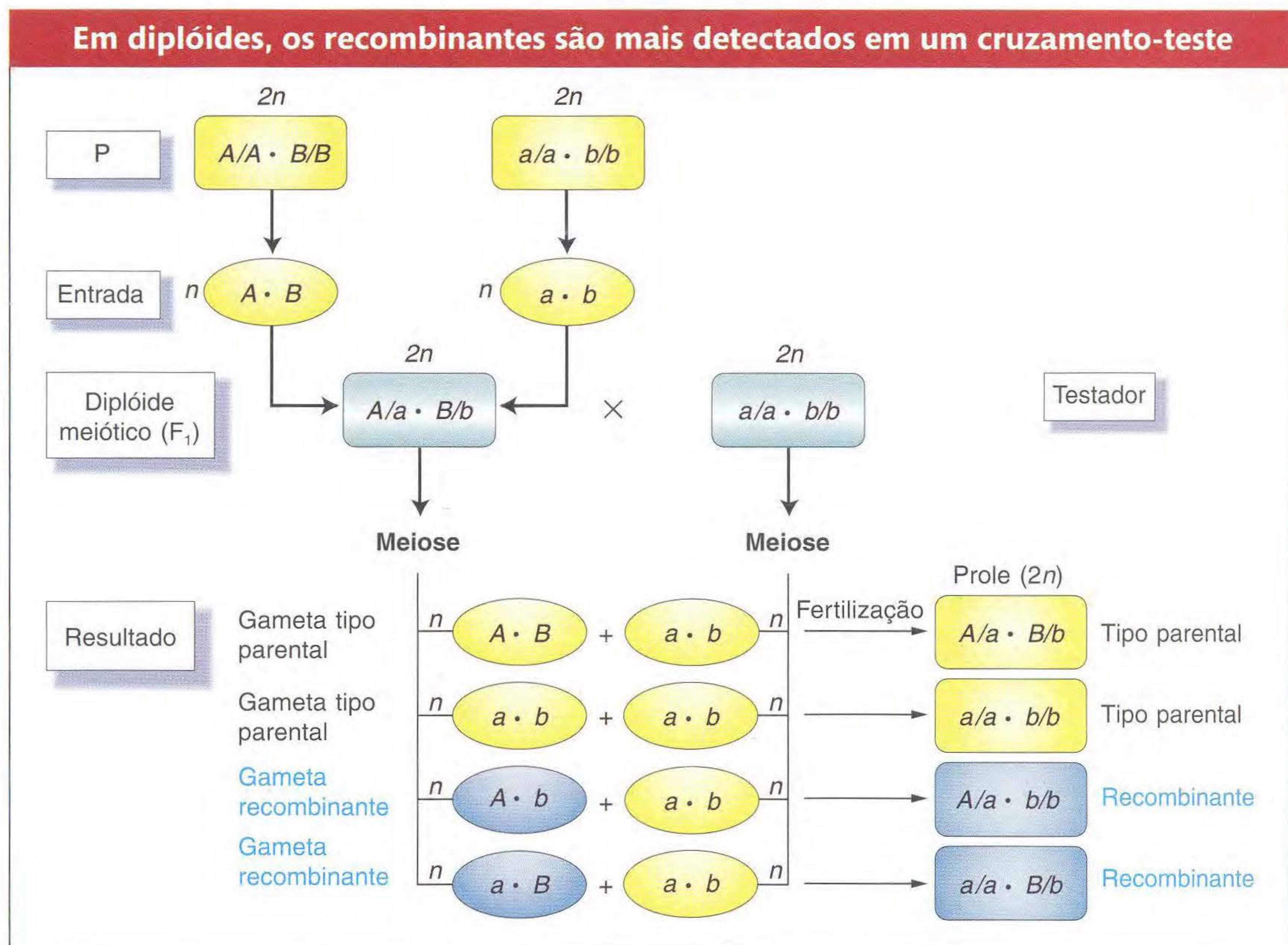


FIG. 3.12 Produtos recombinantes de uma meiose diplóide são mais prontamente detectados em um cruzamento de um heterozigoto e um testador recessivo. Note que a Figura 3.11 é repetida como parte desse diagrama.

distantes no *mesmo* par de cromossomos podem se distribuir praticamente independentemente e produzir o mesmo resultado; veja Capítulo 4.)

Uma frequência de recombinantes de 50% indica que os genes distribuem-se independentemente e, mais provavelmente, estão em cromossomos diferentes.

► 3.4 Herança Poligênica

Até agora, nossa análise neste livro enfocou diferenças monogênicas, com o uso de fenótipos bem contrastantes, tais como pétalas vermelhas *versus* brancas, sementes lisas *versus* rugosas e asas de *Drosophila* longas *versus* vestigiais. Entretanto, uma grande proporção da variação em populações naturais tem a forma de variação *contínua*, que é tipicamente encontrada em características que podem ter algum valor mensurável entre dois extremos. Altura, peso e intensidade de cor são exemplos de tais *características métricas* ou *quantitativas*. Tipicamente, quando o valor métrico dessas características é plotado contra a frequência em uma população natural, a curva de distribuição tem a forma de um sino (Figura 3.14). A forma de sino é devida ao fato de que os valores médios são os mais comuns, enquanto os valores extremos são raros.

Muitos casos de variação contínua têm uma base puramente experimental, pouco afetados pela genética. Por exem-

plo, uma população de plantas geneticamente homozigotas cultivadas em uma área mostra uma curva em forma de sino quanto à altura, com as menores plantas ao redor das margens e as maiores plantas no meio. Essa variação pode ser explicada apenas por fatores ambientais como a mistura e quantidade de fertilizante aplicado. Entretanto, muitos casos de variação contínua têm uma base genética. A cor da pele humana é um exemplo: todos os graus de pele escura podem ser observados na população de partes diferentes do mundo, e essa variação claramente tem um componente genético. Em tais casos, de alguns a muitos alelos interagem com efeito mais ou menos aditivo. Os genes que interagem tendo uma variação contínua hereditária são chamados de **poligenes** ou **loci de característica quantitativa (QTL)**. (O termo locus de característica quantitativa precisa de uma definição: *quantitativa* é mais ou menos sinônimo de contínua; *característica* é mais ou menos sinônimo de caráter ou propriedade; *locus*, literalmente, significa local em um cromossomo e é mais ou menos sinônimo de gene.) Os poligenes, ou QTL, para a mesma característica são distribuídos pelo genoma. Em muitos casos, eles estão em cromossomos diferentes e apresentam distribuição independente, sendo um tópico deste capítulo. Mostraremos como a herança de vários poligenes heterozigotos (até mesmo apenas dois) podem gerar uma curva de distribuição em forma de sino.

Consideremos um modelo simples que foi originalmente usado para explicar a variação contínua no grau de verme-

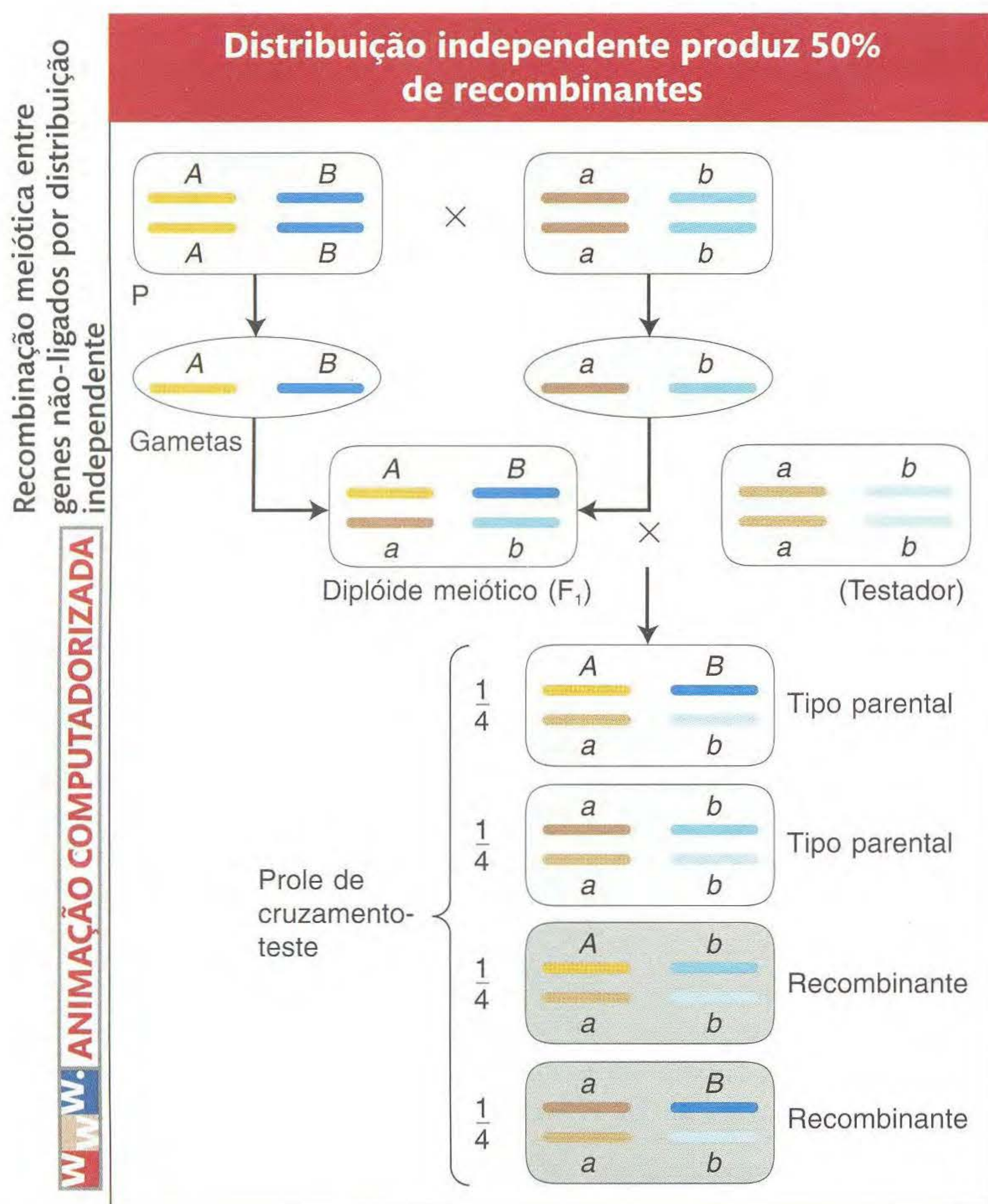


FIG. 3.13 Esse diagrama mostra dois pares de cromossomos de um organismo diplóide com A e a em um par e B e b no outro. A distribuição independente produz uma frequência de recombinantes de 50%. Note que podemos representar a situação haplóide removendo o cruzamento parental (P) e o cruzamento-teste.

lhidão em sementes de trigo. O trabalho foi feito por Hermann Nilsson-Ehle no início do século XX. Vamos supor dois pares de genes de distribuição independente R_1/r_1 e R_2/r_2 . Tanto R_1 quanto R_2 contribuem para a vermelhidão do trigo. Cada “dose” de um alelo R de qualquer dos genes é aditivo, significando que ele aumenta o grau de vermelhidão proporcionalmente. Um cruzamento ilustrativo é a autofecundação de um diíbrido $R_1/r_1; R_2/r_2$. Tanto gametas masculinos quanto femininos apresentarão as seguintes proporções genotípicas:

$\frac{1}{4} R_1; R_2$	2 doses de vermelho
$\frac{1}{4} R_1; r_2$	1 dose de vermelho
$\frac{1}{4} r_1; R_2$	1 dose de vermelho
$\frac{1}{4} r_1; r_2$	0 dose de vermelho

No geral, nessa população de gametas, um quarto tem duas doses, metade tem uma dose, um quarto tem zero doses. A união dos gametas masculino e feminino, ambos mostrando essa composição de R doses, é ilustrada na Figura 3.15. O número de doses na prole varia de quatro ($R_1/R_1; R_2/R_2$) até zero ($r_1/r_1; r_2/r_2$), com todos os valores intercalares.

As proporções na grade da Figura 3.15 podem ser desenhadas como um histograma, como mostrado na Figura 3.16. A forma do histograma pode ser imaginada como um arcabouço que pode ter como base subjacente a curva de distribuição em forma de sino.

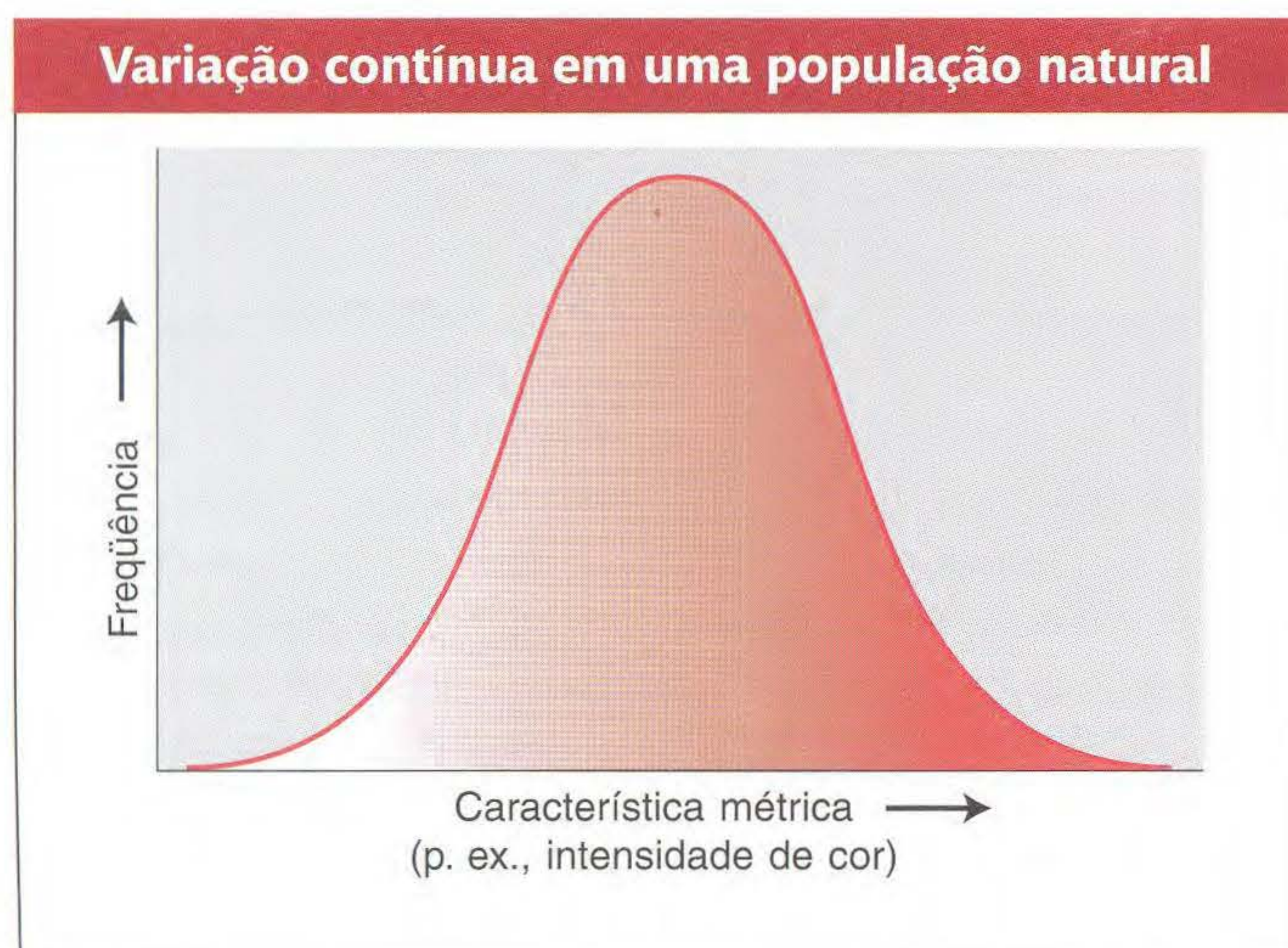


FIG. 3.14 Em uma população, uma característica métrica tal como a intensidade da cor pode ter muitos valores. Assim, a distribuição é na forma de uma curva suave, com os valores mais comuns representando o ponto alto da curva. Se a curva é simétrica, ela é em forma de sino, como mostrado.

Poligenes na prole de uma autofecundação diíbrida

Autofecundação de $R_1/r_1 ; R_2/r_2$

gametas ♂

	2 doses $\frac{1}{4}$	1 dose $\frac{1}{2}$	0 dose $\frac{1}{4}$	
gametas ♀	2 doses $\frac{1}{4}$	4 doses $\frac{1}{16}$	3 doses $\frac{2}{16}$	2 doses $\frac{1}{16}$
	1 dose $\frac{1}{2}$	3 doses $\frac{2}{16}$	2 doses $\frac{4}{16}$	1 dose $\frac{2}{16}$
	0 dose $\frac{1}{4}$	2 doses $\frac{1}{16}$	1 dose $\frac{2}{16}$	0 dose $\frac{1}{16}$

Geral na prole

4 doses $\frac{1}{16}$	3 doses $\frac{4}{16}$	2 doses $\frac{6}{16}$	1 dose $\frac{4}{16}$	0 dose $\frac{1}{16}$
---------------------------	---------------------------	---------------------------	--------------------------	--------------------------

FIG. 3.15 A prole de uma autofecundação diíbrida para dois poligenes pode ser expressa com números de “doses” alélicas aditivas.

Histograma de poligenes de uma autofecundação diíbrida

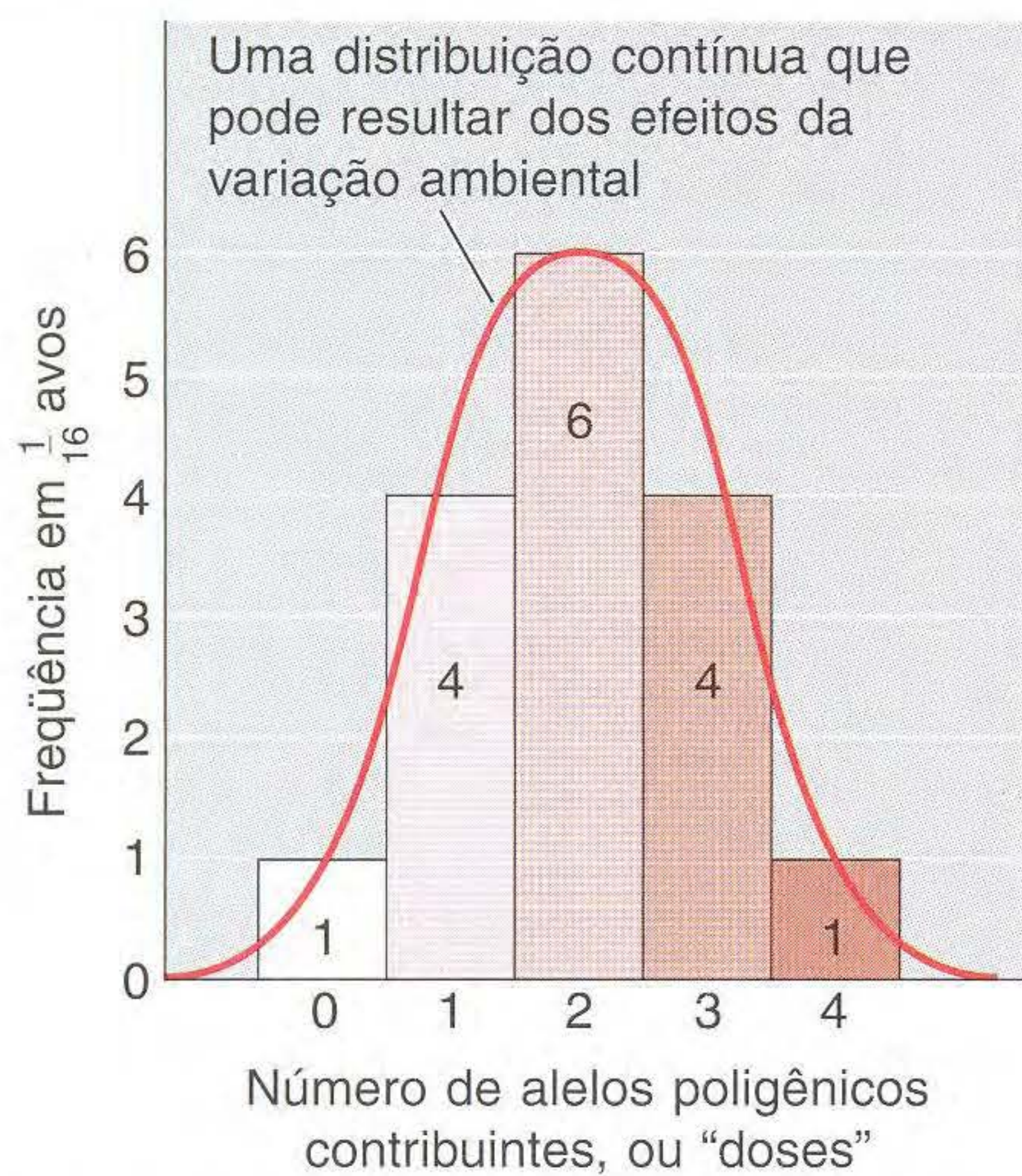


FIG. 3.16 A prole mostrada na Figura 3.15 pode ser representada como um histograma de frequência de alelos poligênicos contribuintes ("doses").

nível de "dose" de um poligene. Supostamente, essa variação dentro de uma classe é o resultado de diferenças ambientais. Assim, o ambiente pode ser visto como contribuindo de um modo que circunda os ombros pontudos do histograma de barras, resultando em uma curva em forma de sino (a linha vermelha no histograma). Se o número de poligenes é aumentado, o histograma aproxima-se mais de uma suave distribuição contínua. Por exemplo, para uma característica determinada por três poligenes, o histograma é como mostrado na Figura 3.17.

Em nossa ilustração, usamos uma autofecundação diíbrida para mostrar como o histograma é produzido. Mas como o nosso exemplo é relevante para o que ocorre em populações naturais? Afinal, nem todos os cruzamentos podem ser desse tipo. Entretanto, se os alelos em cada par de genes são aproximadamente iguais em frequência na população (por exemplo, R_1 é tão comum quanto r_1), o cruzamento diíbrido pode ser dito como representando um cruzamento médio para uma população na qual dois poligenes estão segregando.

A identificação de poligenes e a compreensão de como eles atuam e interagem são desafios importantes para os geneticistas no século XX. A identificação de poligenes será especialmente importante na medicina. Muitas doenças humanas comuns, tais como aterosclerose (endurecimento das artérias) e hipertensão (pressão sanguínea alta), são ditas como tendo um componente poligênico. Assim, a total compreensão dessas condições, que afetam grandes proporções das populações humanas, requer a compreensão desses poligenes, de sua herança e de sua função. Hoje em dia, vários enfoques moleculares podem ser aplicados à tarefa de encontrar poligenes, e consideraremos alguns em capítulos subsequentes. Note que os poligenes não são considerados uma classe funcional especial de genes. Eles são identificados

Histograma de poligenes de uma autofecundação triíbrida

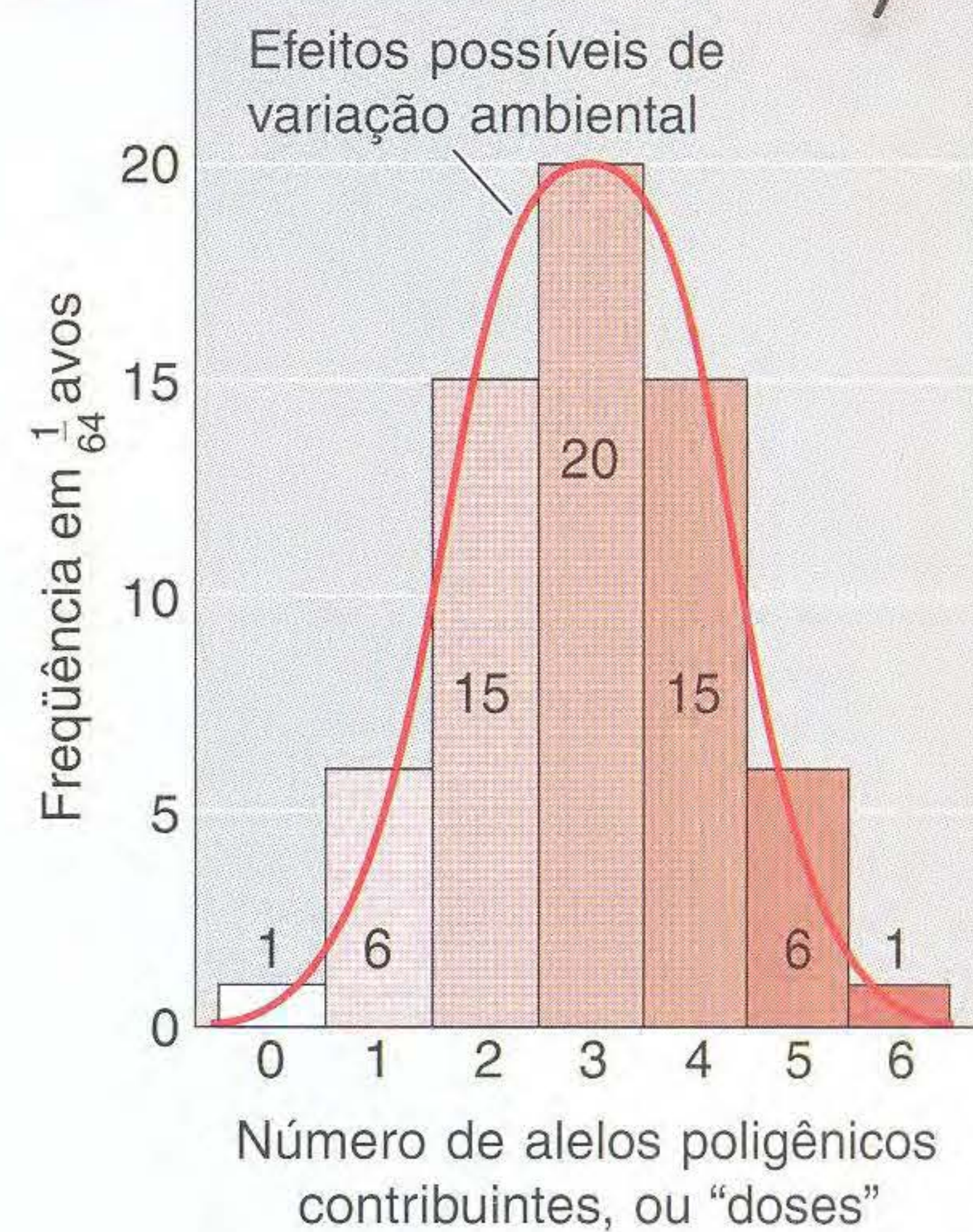


FIG. 3.17 A prole de um triíbrido poligênico pode ser representada por um gráfico de histograma de frequência de alelos poligênicos contribuintes ("doses").

como um grupo apenas no sentido de que têm alelos que contribuem para a variação contínua.

A variação e a distribuição de poligenes podem contribuir para a variação contínua em uma população.

► 3.5 Genes de Organelas: Herança Independente do Núcleo

Até agora consideramos apenas genes nucleares. Embora o núcleo contenha a maioria dos genes de um organismo eucariótico, um subgrupo distinto e especializado do genoma é encontrado nas mitocôndrias, e, nas plantas, também nos cloroplastos. Esses subgrupos são herdados independentemente do genoma nuclear, e assim eles constituem um caso especial de herança independente, às vezes chamada herança extranuclear.

As mitocôndrias e cloroplastos são organelas especializadas situadas no citoplasma. Contêm pequenos cromossomos circulares que levam um subgrupo definido do genoma celular total. Os genes mitocondriais estão envolvidos na tarefa mitocondrial de produção de energia, enquanto os genes de cloroplastos são necessários para o cloroplasto desempenhar sua função de fotossíntese. Entretanto, nenhuma das organelas é funcionalmente autônoma, pois cada uma é baseada em grande parte nos genes nucleares para sua função. A razão pela qual alguns dos genes necessários estão nas próprias organelas e outros no núcleo ainda é um certo mistério, que não será abordado aqui.

Outra peculiaridade dos genes de organelas é o grande número de cópias presentes em uma célula. Cada organela

Células mostrando nucleóides dentro de mitocôndria



FIG. 3.18 Coloração fluorescente de uma célula de *Euglena gracilis*. Com os corantes usados, o núcleo aparece em vermelho devido à fluorescência de grandes quantidades de DNA nuclear. As mitocôndrias fluorescem em verde, e, dentro das mitocôndrias, as concentrações de DNA mitocondrial (nucleóides) fluorescem em amarelo. [De Y. Huyashi and K. Veda, J. Cell Sci. 93, 1989, 565.]

está presente em muitas cópias por célula, e, além disso, cada organela contém muitas cópias de seu cromossomo. Assim, cada célula pode conter centenas ou milhares de cromossomos de organela. Considere os cloroplastos, por exemplo. Qualquer célula verde de uma planta tem muitos cloroplastos, e cada cloroplasto contém muitas moléculas circulares de DNA idênticas, os chamados cromossomos de cloroplastos. Assim, o número de cromossomos de cloroplasto por célula pode ser de milhares, e o número ainda pode variar de célula a célula. O DNA às vezes é visto embalado em estruturas suborganelares chamadas *nucleóides*, que se tornam visíveis se corados com um corante de ligação ao DNA (Figura 3.18). O DNA é dobrado dentro do nucleóide, mas não tem o tipo de helicoidização com histonas associadas mostrado por cromossomos nucleares. O mesmo arranjo é verdadeiro para o DNA nas mitocôndrias. No momento, devemos supor que todas as cópias de um cromossomo de organela dentro de uma célula são idênticas, mas teremos que afrouxar essa suposição depois.

Muitos cromossomos de organelas hoje foram seqüenciados. Os exemplos do tamanho relativo e espaçamento no DNA mitocondrial (mtDNA) e DNA de cloroplasto (cpDNA) são mostrados na Figura 3.19. Os genes de organelas são proximamente espaçados, e, em alguns organismos, podem conter íntrons. Note como a maioria dos genes envolvem

Genomas de organela

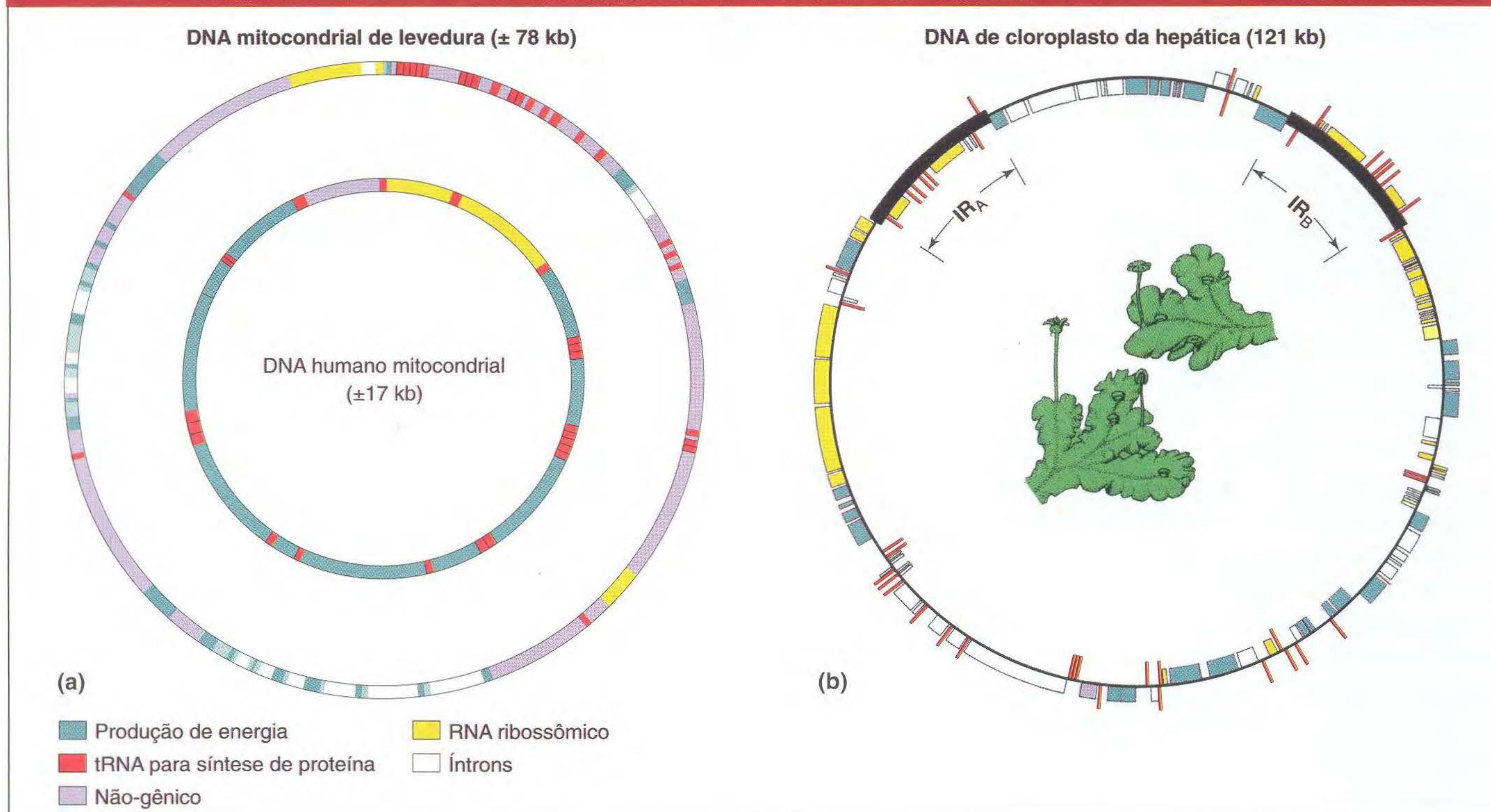


FIG. 3.19 Mapas de DNA para mitocôndrias e cloroplastos. Muitos dos genes de organelas codificam proteínas que fazem as funções de produção de energia dessas organelas (verde), enquanto outras (vermelha e laranja) funcionam na síntese de proteínas. (a) Mapas de levedura e mtDNA humanos. (Note que o mapa humano não é desenhado na mesma escala que o mapa de levedura.) (b) O genoma de cloroplasto com 121 kb da hepática *Marchantia polymorpha*. Os genes mostrados dentro do mapa são transcritos no sentido horário, e os mostrados fora são transcritos no sentido anti-horário. IR_A e IR_B indicam repetições invertidas. O desenho superior no centro do mapa representa uma planta macho de *Marchantia*; o desenho inferior representa uma fêmea. [De K. Umesono and H. Ozeki, Trends Genet. 3, 1987.]

reações químicas que ocorrem dentro da própria organela: a fotossíntese nos cloroplastos e a fosforilação oxidativa nas mitocôndrias.

Padrões de herança em organelas

Os genes de organelas apresentam seu próprio modo especial de herança chamada de herança uniparental: a prole herda genes de organela exclusivamente de um genitor, mas não do outro. Na maioria dos casos, esse genitor é a mãe, um padrão chamado de **herança materna**. Por que só a mãe? A resposta está no fato de os cromossomos de organelas estarem situados no citoplasma e os gametas femininos não contribuírem com citoplasma igualmente para o zigoto. Com relação aos genes nucleares, ambos os genitores contribuem igualmente para o zigoto. Entretanto, o ovócito contribui com a maior parte do citoplasma, enquanto o espermatozóide não contribui com praticamente nada. Assim, como as organelas estão no citoplasma, o genitor feminino contribui com as organelas ao longo do citoplasma, e essencialmente nada do DNA da organela no zigoto vem do genitor masculino.

Algumas variantes fenotípicas são causadas por um alelo mutante de um gene de organela, e podemos usar esses mutantes para rastrear padrões da herança de organelas. Vamos temporariamente supor que o alelo mutante está presente em todas as cópias do cromossomo da organela, uma situação que é geralmente encontrada. Em um cruzamento, o fenótipo variante será transmitido para a prole se a variante usada é o genitor feminino, mas não se é o genitor masculino. Assim, em geral, a herança citoplasmática mostra o seguinte padrão:

mutante feminino \times macho tipo selvagem \rightarrow prole toda mutante
fêmea tipo selvagem \times macho mutante \rightarrow prole toda tipo selvagem

De fato, esse padrão de herança é diagnóstico da herança de organelas em casos nos quais a localização genômica de um alelo mutante não é conhecida.

A herança materna pode ser claramente demonstrada em alguns mutantes de fungos. Por exemplo, no fungo *Neurospora*, um mutante chamado *poky* tem um fenótipo de crescimento lento. A *Neurospora* pode ser cruzada de tal modo que um genitor atua como genitor materno, contribuindo com o citoplasma (veja Figura 3.9). Os resultados dos seguintes cruzamentos recíprocos sugerem que o gene mutante reside nas mitocôndrias (fungos não têm cloroplastos):

fêmea *poky* \times macho tipo selvagem \rightarrow prole toda *poky*
fêmea tipo selvagem \times macho *poky* \rightarrow prole toda tipo selvagem

O sequenciamento mostrou que o fenótipo *poky* é causado por uma mutação de um gene de RNA ribossômico no mtDNA. Sua herança é mostrada diagramaticamente na Figura 3.20. O cruzamento inclui uma diferença alélica (*ad* e *ad*⁺) em um gene nuclear em adição a *poky*; note como a herança mendeliana do gene nuclear é independente da herança materna do fenótipo *poky*.

Fenótipos variantes causados por mutações no DNA de organela citoplasmática em geral são herdados maternamente e não dependem dos padrões mendelianos apresentados pelos genes nucleares.

Segregação citoplasmática

Em alguns casos, as células contêm misturas de organelas mutantes e normais. Essas células são chamadas de *citohets* ou *heteroplasmons*. Nessas misturas, um tipo de **segregação**

Herança materna de mutante mitocondrial de fenótipo poky

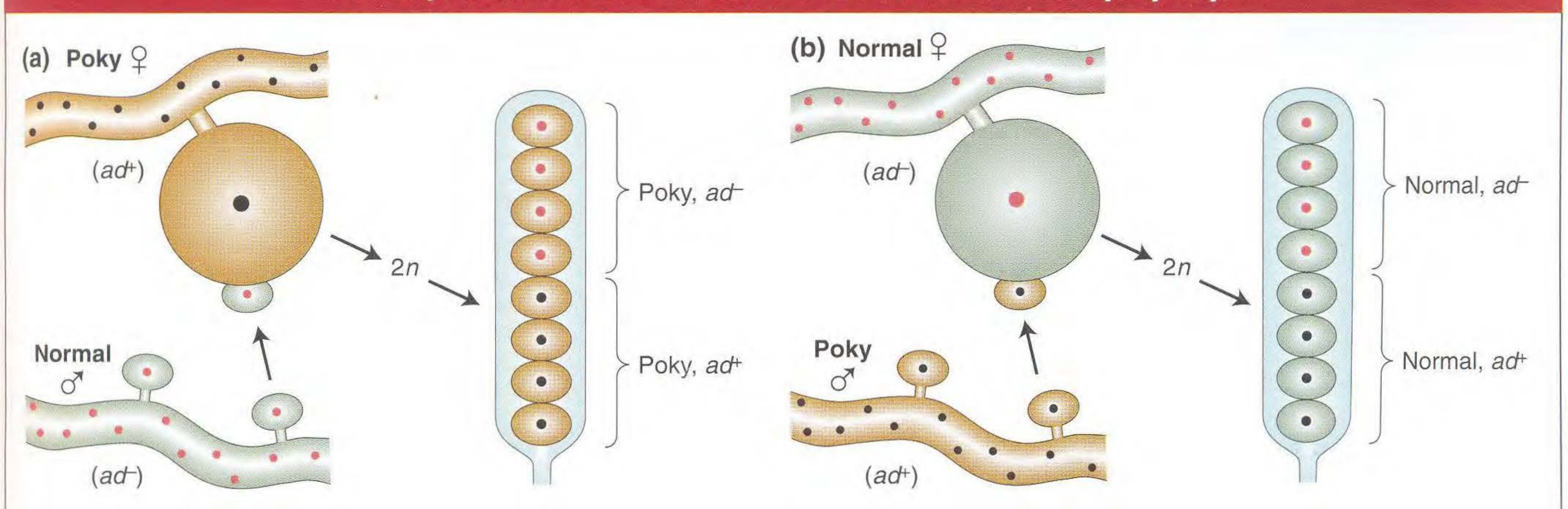


FIG. 3.20 Cruzamentos recíprocos de *poky* e *Neurospora* tipo selvagem produzem resultados diferentes porque um genitor diferente contribui com o citoplasma. O genitor feminino contribui com a maioria do citoplasma das células-filhas. O sombreado marrom representa o citoplasma com mitocôndrias contendo a mutação *poky*, e o sombreado verde representa citoplasma com mitocôndrias tipo selvagem. Note que toda a prole na parte a é *poky*, enquanto toda a prole na parte b é normal. Assim, ambos os cruzamentos apresentam herança materna. O gene nuclear com os alelos *ad*⁺ (preto) e *ad*⁻ (vermelho) são usados para ilustrar a segregação dos genes nucleares na proporção mendeliana 1:1 esperada para esse organismo haplóide.

citoplasmática pode ser detectada, na qual os dois tipos se distribuem em células-filhas diferentes. O processo mais provavelmente é baseado na repartição aleatória no curso da divisão celular. As plantas fornecem um bom exemplo. Muitos casos de folhas brancas são causados por mutações nos genes de cloroplastos que controlam a produção e deposição do pigmento verde clorofila. Como a clorofila é necessária para a planta viver, esse tipo de mutação é letal, e as plantas de folhas brancas não podem ser obtidas de cruzamentos experimentais. Entretanto, algumas plantas são variegadas, tendo tanto manchas verdes quanto brancas. Assim, as plantas variegadas fornecem um meio para demonstrar a segregação citoplasmática.

A planta maravilha na Figura 3.21 mostra uma folha variegada comumente observada e o fenótipo de uma ramificação que demonstra a herança de um alelo mutante de um gene de cloroplasto. O alelo mutante faz com que os cloroplastos sejam brancos; por sua vez, a cor dos cloroplastos determina a cor das células e, portanto, a cor dos ramos compostos dessas células. Os ramos variegados são mosaicos de células todas brancas e células todas verdes. As flores podem se desenvolver em ramos verdes, brancos ou variegados, e os genes de cloroplastos das células de uma flor são os ramos nos quais elas crescem. Assim, em um cruzamento (Figura 3.22), o gameta materno dentro da flor (o ovócito) determina

Folhas variegadas causadas por uma mutação em cpDNA

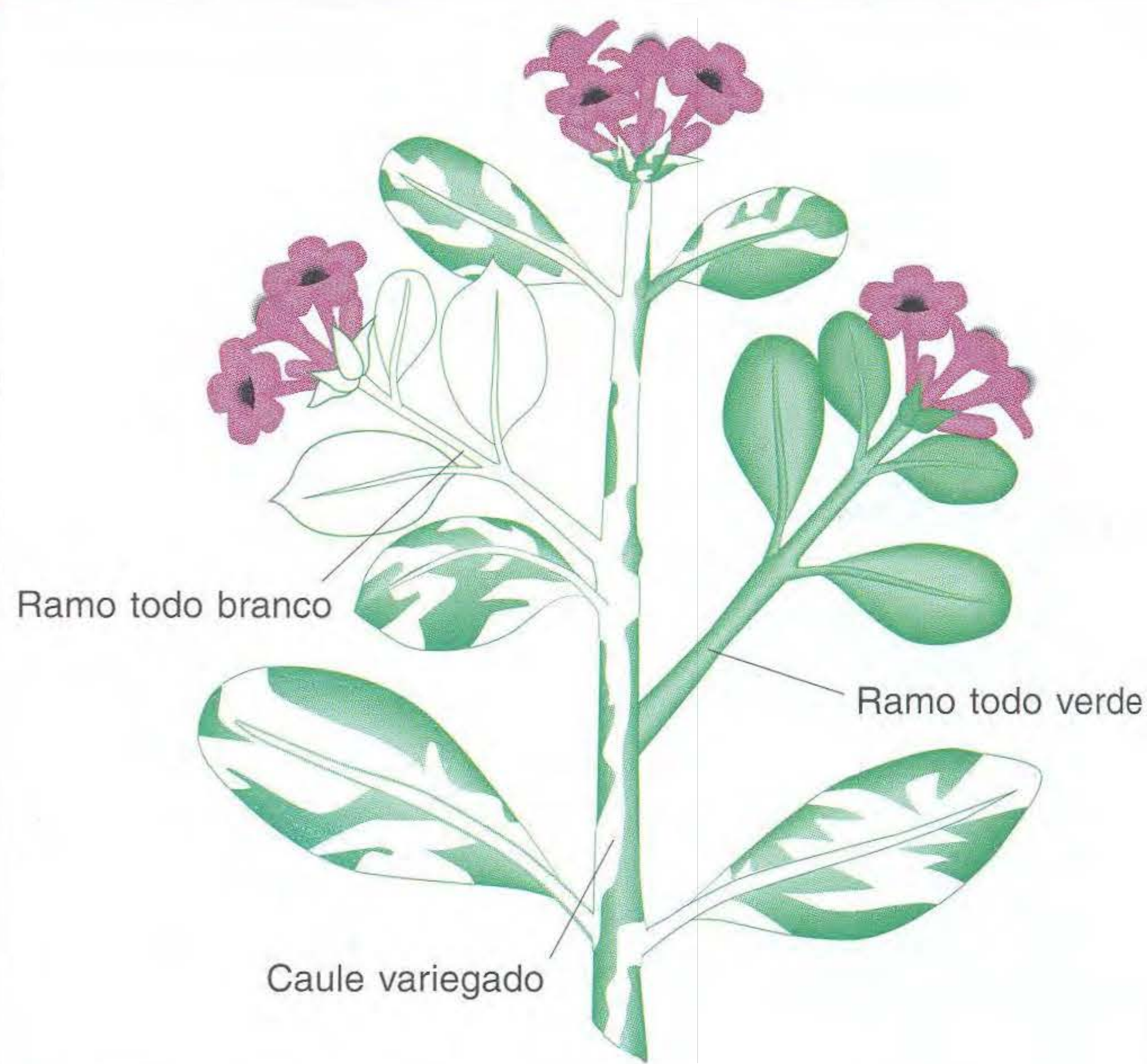


FIG. 3.21 Variação de folha em *Mirabilis jalapa*, a planta maravilha. As flores podem ser formadas em qualquer ramo (variegadas, verdes ou brancas), e essas flores podem ser usadas em cruzamentos.

Cruzamentos usando flores de uma planta variegada

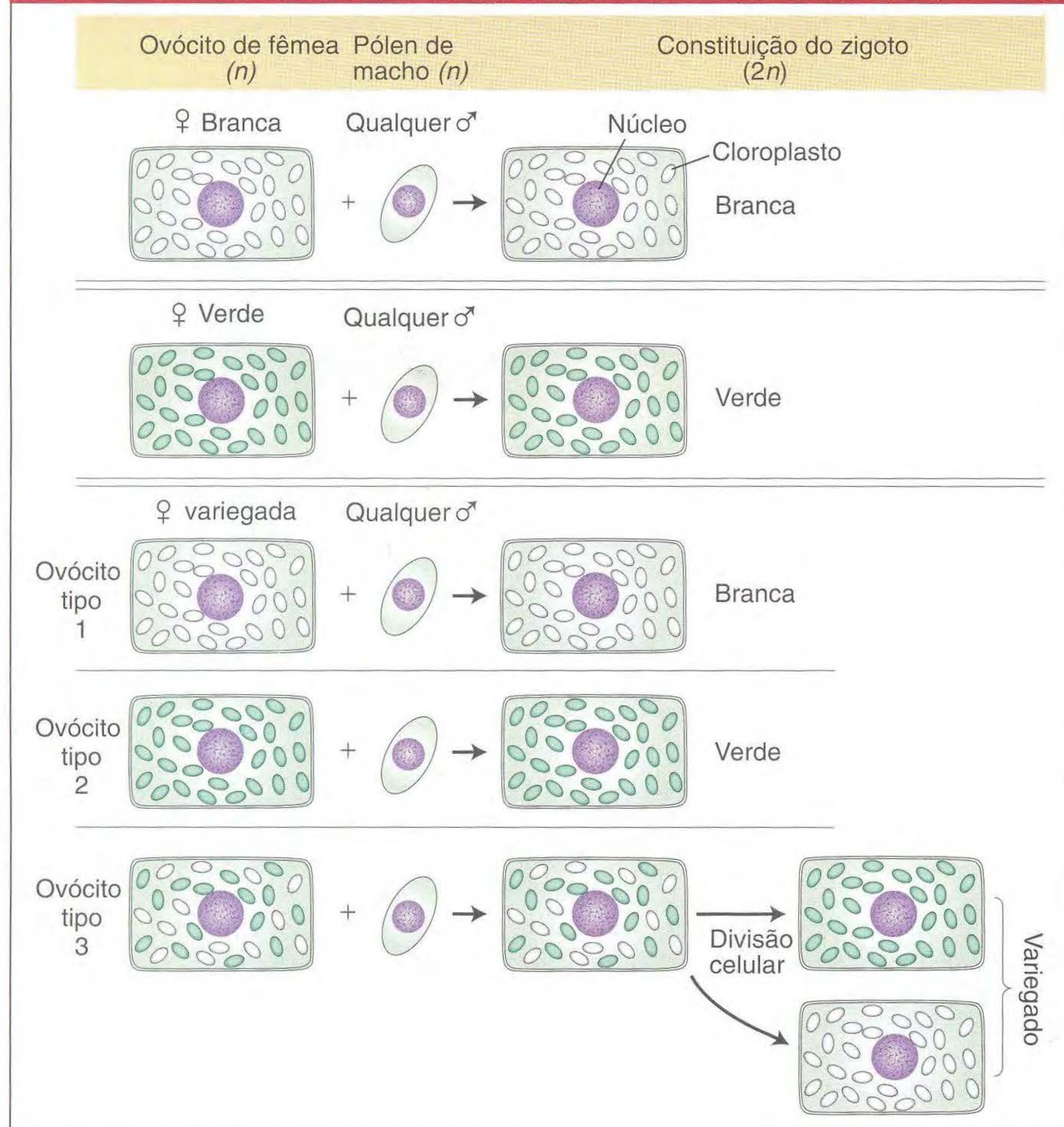


FIG. 3.22 Os resultados dos cruzamentos de *Mirabilis jalapa* podem ser explicados por herança autônoma de cloroplastos. As grandes esferas escuras representam núcleos. Os corpos menores representam cloroplastos, verdes ou brancos. Cada ovócito supostamente contém muitos cloroplastos, e cada célula de pólen supostamente não contém cloroplastos. Os primeiros dois cruzamentos exibem estrita herança materna. Se, entretanto, o ramo materno é variegado, podem resultar três tipos de zigotos, dependendo de o ovócito conter apenas cloroplastos brancos, apenas verdes ou verdes e brancos. No último caso, o zigoto resultante pode produzir tanto tecido verde quanto branco, e assim resulta uma planta variegada.

a cor das folhas e ramos da planta prole. Por exemplo, se um ovócito é de uma flor em um ramo verde, toda a prole será verde, independentemente da origem do pólen. Um ramo branco terá cloroplastos brancos, e a prole resultante dessas plantas será branca. (Devido à letalidade, os descendentes brancos não vivem além do estágio de semente.)

Os zigotos variegados (parte inferior da Figura 3.22) demonstram segregação citoplasmática. Essa prole variegada vem de ovócitos que são citohets. Curiosamente, quando tal zigoto se divide, os cloroplastos brancos e verdes em geral se segregam; isto é, eles se distribuem em células separadas, produzindo os setores verde e branco que causam a variegação nos ramos. Aqui está uma demonstração direta de segregação citoplasmática.

Considerando que uma célula é uma população de moléculas de organelas, como é possível obter uma célula mutante “pura”, contendo apenas cromossomos mutantes? Mais provavelmente, mutantes puros são criados em células assexuais como segue. As variantes surgem por mutação de um único gene em um único cromossomo. Então, em alguns casos, o cromossomo tendo a mutação pode por acaso aumentar em frequência na população dentro da célula. Esse processo é chamado de *deriva genética aleatória*. Uma célula que é citohet tem, digamos, 60% de cromossomos *A* e 40% de cromossomos *a*. Quando essa célula se divide, às vezes todos os cromossomos *A* vão para uma célula-filha, e todos os cromossomos *a* para outra (novamente ao acaso). Mais frequentemente, essa repartição requer várias gerações subsequentes de divisão celular para se completar (Figura 3.23). Assim, como resultado desses eventos casuais, ambos os alelos são expressos em células-filhas

diferentes, e essa separação continuará nas descendentes dessas células. Note que a segregação citoplasmática não é um processo mitótico; ela ocorre em células assexuais em divisão, mas não está relacionada à mitose. Nos cloroplastos, a segregação citoplasmática é um mecanismo comum para produzir plantas variegadas (verde e branca), como já mencionado. Em mutantes de fungos tais como o mutante *poky* de *Neurospora*, a mutação original em uma molécula de mtDNA deve ter se acumulado e sofrido segregação citoplasmática para produzir a linhagem que expressa os sintomas *poky*.

As populações de organelas que contêm misturas de dois cromossomos geneticamente distintos em geral apresentam segregação de dois tipos nas células-filhas na divisão celular. Esse processo é chamado de segregação citoplasmática.

Em alguns sistemas especiais tais com fungos e algas, os citohets que são “diíbridos” foram obtidos (digamos, *AB* em um cromossomo de organela e *ab* em outro). Em tais casos, raros processos tipo crossing podem ocorrer, mas tal ocorrência deve ser considerada um fenômeno genético menor.

Alelos em cromossomos de organelas

1. Em cruzamentos sexuais são herdados só de um genitor (geralmente o materno) e, assim, não mostram proporções de segregação do tipo que os genes nucleares fazem.
2. Em ciclos assexuais podem apresentar segregação citoplasmática.
3. Em células assexuais podem ocasionalmente apresentar processos análogos ao crossing over.

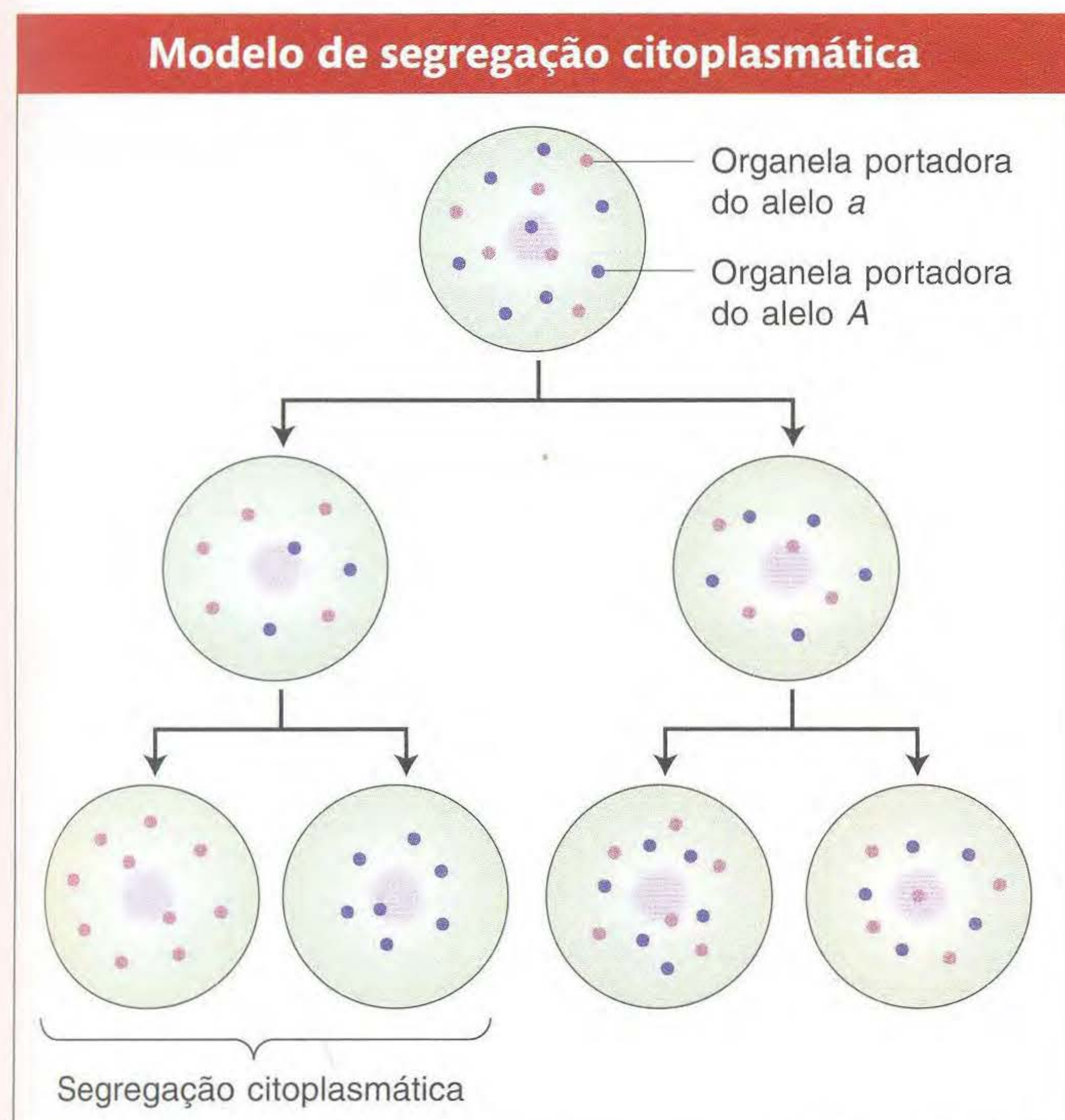


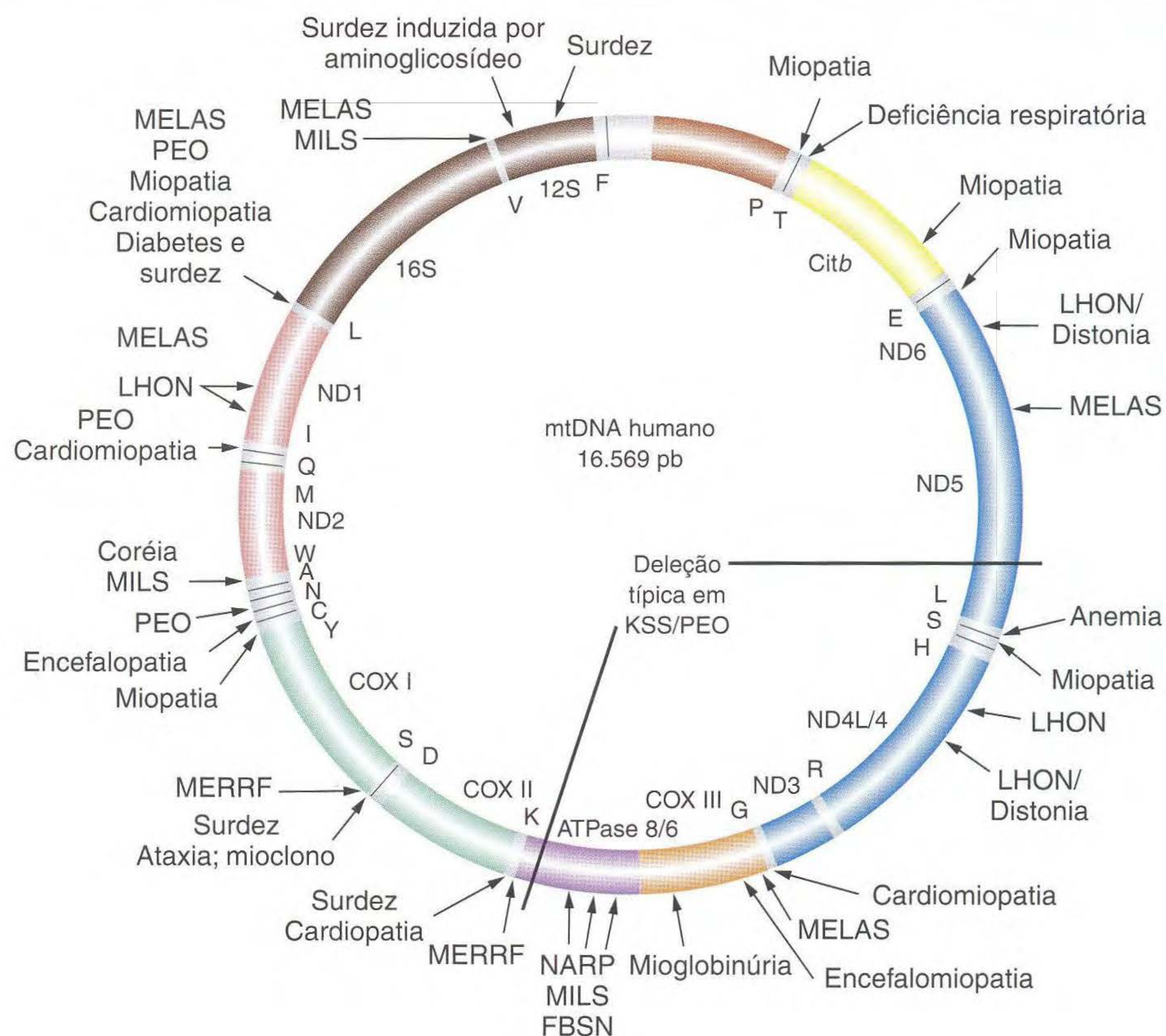
FIG. 3.23 Ao acaso, organelas geneticamente distintas podem segregar em células separadas em várias divisões celulares sucessivas. Os pontos vermelhos e azuis representam organelas geneticamente distinguíveis, tais como mitocôndrias com e sem uma mutação.

Mutações citoplasmáticas em humanos

Existem mutações citoplasmáticas em humanos? Alguns heredogramas humanos apresentam a transmissão de raros distúrbios apenas através de mulheres e nunca através de homens. Esse padrão sugere fortemente a herança citoplasmática e indica uma mutação no mtDNA como o motivo para o fenótipo. A doença MERRF (epilepsia mioclônica e fibras vermelhas anfractuadas) é um desses fenótipos, resultando da mudança de uma única base no mtDNA. É uma doença que afeta os músculos, mas os sintomas também incluem distúrbios oculares e auditivos. Um outro exemplo é a síndrome de Kearns-Sayre, uma constelação de sintomas afetando os olhos, coração, músculos e cérebro que é causada pela perda de uma parte do mtDNA. Em alguns desses casos, as células de um paciente contêm misturas de cromossomos normais e mutantes, e as proporções de cada uma passada para a prole podem variar como resultado de segregação citoplasmática. As proporções em uma pessoa também podem variar em tecidos diferentes ou com o tempo. O acúmulo de alguns tipos de mutações mitocondriais com o tempo tem sido proposto como uma possível causa de envelhecimento.

A Figura 3.24 mostra algumas mutações em genes mitocondriais humanos que podem levar a uma doença quando, por deriva aleatória e segregação citoplasmática, eles aumentam de frequência de tal modo que o funcionamento celular é prejudicado. A herança de uma doença mitocondrial humana é mostrada na Figura 3.25. Note que a condição é

Sítios de mutações no mtDNA em algumas doenças humanas



Doenças:

MERRF	Epilepsia mioclônica e fibras vermelhas anfractuadas
LHON	Neuropatia óptica hereditária
NARP	Fraqueza muscular neurogênica, ataxia e retinite pigmentosa
MELAS	Encefalomiopatia mitocondrial, acidose láctica e sintomas tipo ataque
MMC	Miopatia e cardiomiopatia herdadas maternamente
PEO	Oftalmoplegia externa progressiva
KSS	Síndrome de Kearns-Sayre
MILS	Síndrome de Leigh herdada maternamente

FIG. 3.24 Esse mapa de mtDNA humano mostra loci de mutações que levam a citopatias. Os genes para RNA transportador são representados por abreviações de aminoácidos por uma só letra; ND = NADH desidrogenase; COX = citocromo oxidase; e 12S e 16S referem-se a RNA ribossômicos. [De S. DiMauro et al., "Mitochondria in Neuromuscular Disorders," *Biochim. Biophys. Acta* 1366, 1998, 199-210.]

Heredograma de uma doença mitocondrial

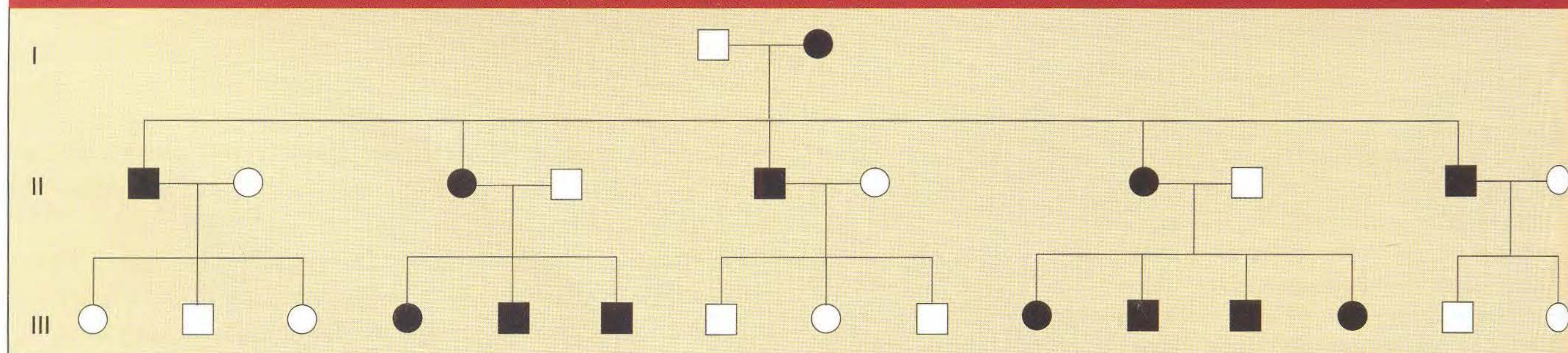


FIG. 3.25 Esse heredograma mostra que uma doença mitocondrial humana é herdada apenas da mãe.

sempre passada para a prole pelas mães e nunca pelos pais. Ocasionalmente, uma mãe produzirá uma criança não afe-

tada (não mostrada), provavelmente devido à segregação citoplasmática no tecido formador da gametas.

Resumo

A pesquisa genética e os cruzamentos de plantas e animais geralmente necessitam da síntese de genótipos que são combinações complexas de alelos de genes diferentes. Tais genes podem estar no mesmo cromossomo ou em cromossomos diferentes; este último é o assunto principal deste capítulo.

No caso mais simples, um diíbrido para o qual os dois pares de genes estão em pares diferentes de cromossomos, cada gene individual mostra segregação igual na meiose como previsto pela Primeira Lei de Mendel. Como as fibras do fuso nuclear se ligam aleatoriamente aos centrômeros na meiose, os dois pares de genes são repartidos independentemente nos produtos meióticos. Esse princípio de distribuição independente é chamado Segunda Lei de Mendel porque Mendel foi o primeiro a observá-lo. De um diíbrido $A/a ; B/b$ são produzidos quatro genótipos de produtos meióticos: $A ; B ; A ; b ; a ; B$ e $a ; b$ com igual frequência de 25% cada. Assim, em um cruzamento-teste de um diíbrido com um duplo recessivo, as proporções fenotípicas da prole também são de 25% (uma proporção 1:1:1:1). Se tal diíbrido for autofecundado, as classes fenotípicas na prole são $\frac{9}{16} A/- ; B/- ; \frac{3}{16} A/- ; b/b ; \frac{3}{16} a/a ; B/-$ e $\frac{1}{16} a/a ; b/b$. As proporções 1:1:1:1 e 9:3:3:1 são ambas diagnósticas de distribuição independente.

Genótipos mais complexos compostos de genes que se distribuem independentemente podem ser tratados como extensões do caso de segregação monogênica. As proporções gerais genotípicas, fenotípicas ou gaméticas são calculadas aplicando a regra do produto, isto é, multiplicando as proporções relevantes de genes individuais. A probabilidade de ocorrência de qualquer das várias categorias de prole é calculada aplicando a regra da soma, isto é, adicionando suas probabilidades individuais. De forma mnemônica, a regra do produto lida com “A e B,” enquanto a regra da soma lida com “A’ ou A’”. O teste do χ^2 pode ser usado para testar se as proporções observadas de classes na análise genética estão de acordo com as expectativas de uma hipótese genética, tais como a hipótese de herança de um ou dois genes. Se o valor de probabilidade de menos de 5% for calculado, a hipótese deve ser rejeitada.

Gerações seqüenciais de autofecundação aumentam as proporções de homozigotos, de acordo com os princípios de segregação igual e distribuição independente (se os genes

estão em cromossomos diferentes). Assim, a autofecundação é usada para criar linhagens puras complexas com combinações de mutações desejáveis.

A distribuição independente de cromossomos na meiose pode ser observada citologicamente usando pares cromossômicos heteromórficos (os que apresentam uma diferença estrutural). Os cromossomos X e Y são um desses casos, mas outros casos mais raros podem ser encontrados e usados para essa demonstração. A distribuição independente dos genes a nível de meiócitos únicos pode ser observada nos fungos ascomicetos, pois os ascos apresentam os dois tipos alternativos de segregações com frequências iguais.

Uma das principais funções da meiose é produzir recombinantes, novas combinações de alelos de genótipos haplóides que se unem para formar o meiócito. A distribuição independente é a principal fonte de recombinantes. Em um cruzamento-teste diíbrido mostrando distribuição independente, a frequência de recombinantes será de 50%.

Características métricas tais como a intensidade da cor mostram uma distribuição contínua em uma população. Distribuições contínuas podem ser baseadas em variação ambiental ou em alelos variantes de múltiplos genes, ou em uma combinação de ambos. Um simples modelo genético propõe que os alelos ativos de vários genes (chamados de poligenes) contribuem mais ou menos aditivamente para uma característica métrica. Em uma análise da prole da autofecundação ou uma multiplicação de um indivíduo heterozigoto, o histograma mostrando a proporção de cada genótipo aproxima-se de uma curva em forma de sino típica da variação contínua.

Os pequenos subgrupos do genoma encontrados nas mitocôndrias e cloroplastos são herdados independentemente do genoma nuclear. Mutantes nesses genes de organelas geralmente apresentam herança materna, juntamente com o citoplasma, que é a localização dessas organelas. Em citoplasmas geneticamente mistos (citohets), os dois genótipos (digamos, tipo selvagem e mutante) em geral distribuem-se em células-filhas diferentes por um processo pouco compreendido chamado segregação citoplasmática. A mutação mitocondrial em humanos resulta em doenças que apresentam segregação citoplasmática em tecidos do corpo e herança materna em uma reprodução.

Termos-chave

cruzamento diíbrido	locus de característica quantitativa (QTL)	regra da soma
diíbrido	poligene (locus de característica	regra do produto
distribuição independente	quantitativa)	segregação citoplasmática
DNA de cloroplasto (cpDNA)	recombinação	Segunda Lei de Mendel
DNA mitocondrial (mtDNA)	recombinação meiótica	teste do qui-quadrado (χ^2)
herança materna	recombinante	vigor híbrido

► Problemas Resolvidos

Problema resolvido 1. Duas moscas *Drosophila* que tinham asas normais (transparentes, longas) foram cruzadas. Na prole, surgiram dois novos fenótipos: asas opacas (tendo um aspecto semi-opaco) e asas aparadas (com pontas quadradas). A prole foi a seguinte:

Fêmeas	Machos
179 transparentes, longas	92 transparentes, longas
58 transparentes, aparadas	89 opacas, longas
	28 transparentes, aparadas
	31 opacas, aparadas

- Forneça uma explicação cromossômica para esses resultados, mostrando os genótipos cromossômicos dos genitores e de todas as classes de prole sob seu modelo.
- Projete um teste para seu modelo.

SOLUÇÃO

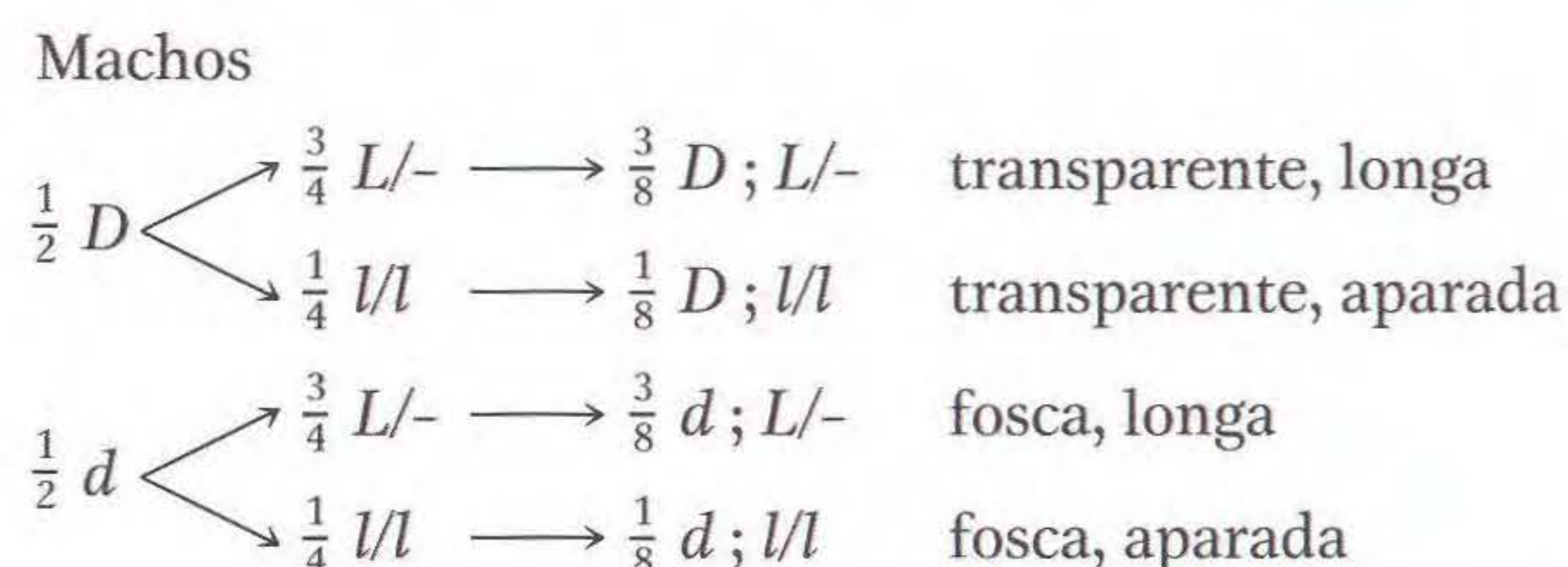
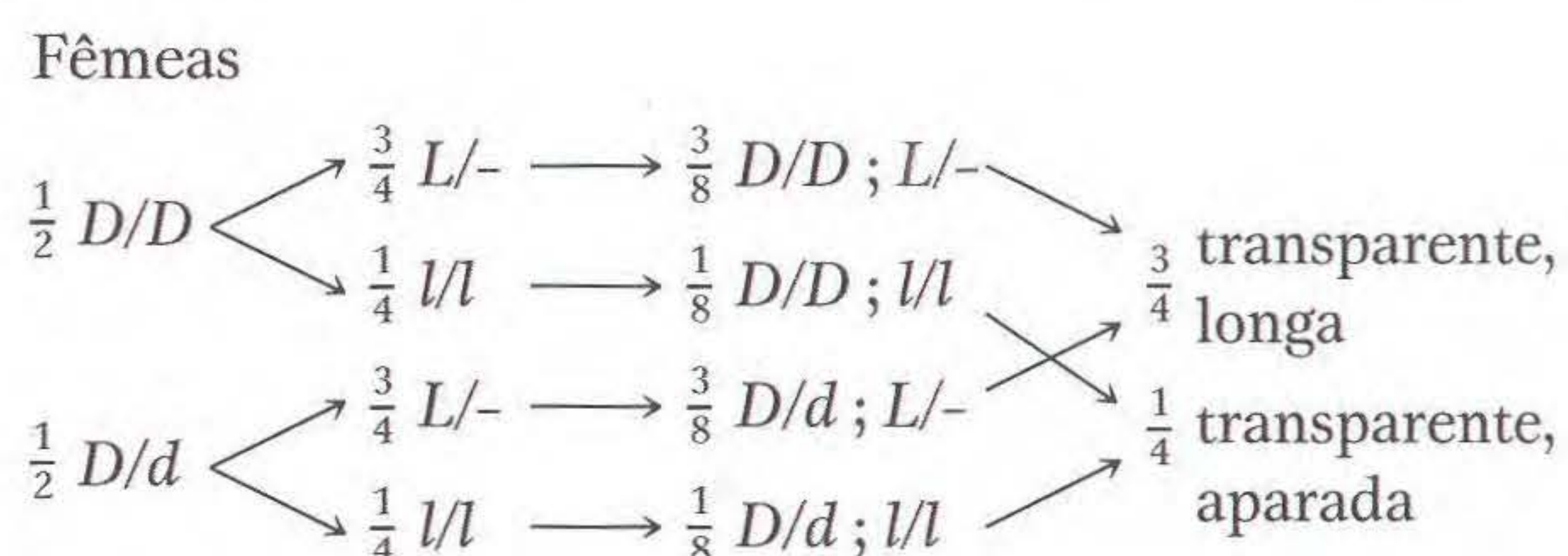
a. A primeira etapa é perceber quaisquer características interessantes dos dados. A primeira característica marcante é o aparecimento de dois novos fenótipos. Encontramos o fenômeno no Capítulo 2, onde foi explicado como os alelos recessivos mascaram suas contrapartes dominantes. Assim, primeiro, devemos supor que uma ou ambas as moscas parentais têm alelos recessivos de dois genes diferentes. Essa interferência é fortalecida pela observação de que alguma prole só expressa um dos novos fenótipos. Se os novos fenótipos sempre apareciam juntos, podemos supor que o mesmo alelo recessivo determina ambos.

Entretanto, a outra característica marcante dos dados, que não podemos explicar usando os princípios mendelianos do Capítulo 2, é a diferença óbvia entre os sexos; embora existam números aproximadamente iguais de machos e fêmeas, os machos caem em quatro classes fenotípicas, mas as fêmeas constituem apenas duas. Esse fato deve sugerir imediatamente algum tipo de herança ligada ao sexo. Quando estudamos os dados, vemos que os fenótipos longa e aparada estão segregando tanto em machos quanto em fêmeas, mas apenas os machos têm o fenótipo opaco. Essa observação sugere que a herança da transparência da asa difere da herança da forma da asa. Primeiro, longa e aparada são encontradas em uma proporção de 3:1 tanto em machos quanto em fêmeas. Essa proporção pode ser explicada se ambos os genitores forem heterozigotos para um gene autossômico; podemos representá-las como L/l , onde L representa longa e l representa aparada.

Tendo feito essa análise parcial, vemos que apenas a herança da transparência da asa está associada ao sexo. A possibilidade mais óbvia é que os alelos para transparente (D) e opaca (d) estão no cromossomo X, pois vimos, no Capítulo 2, que a localização do gene nesse cromossomo dá padrões de herança correlacionados ao sexo. Se essa sugestão é verdadeira, então a fêmea parental deve ser a que tem o alelo d , pois, se o macho tivesse o d , ele seria opaco, enquanto sabemos que ele tinha asas transparentes. Portanto, o genitor feminino seria D/d e o macho D . Veja-

mos se essa sugestão funciona: se for verdadeira, toda a prole feminina herdaria o alelo D de seu pai, e, assim, teria asas transparentes, como foi observado. Metade dos filhos seriam D (asas transparentes) e metade d (asas opacas), o que também foi observado.

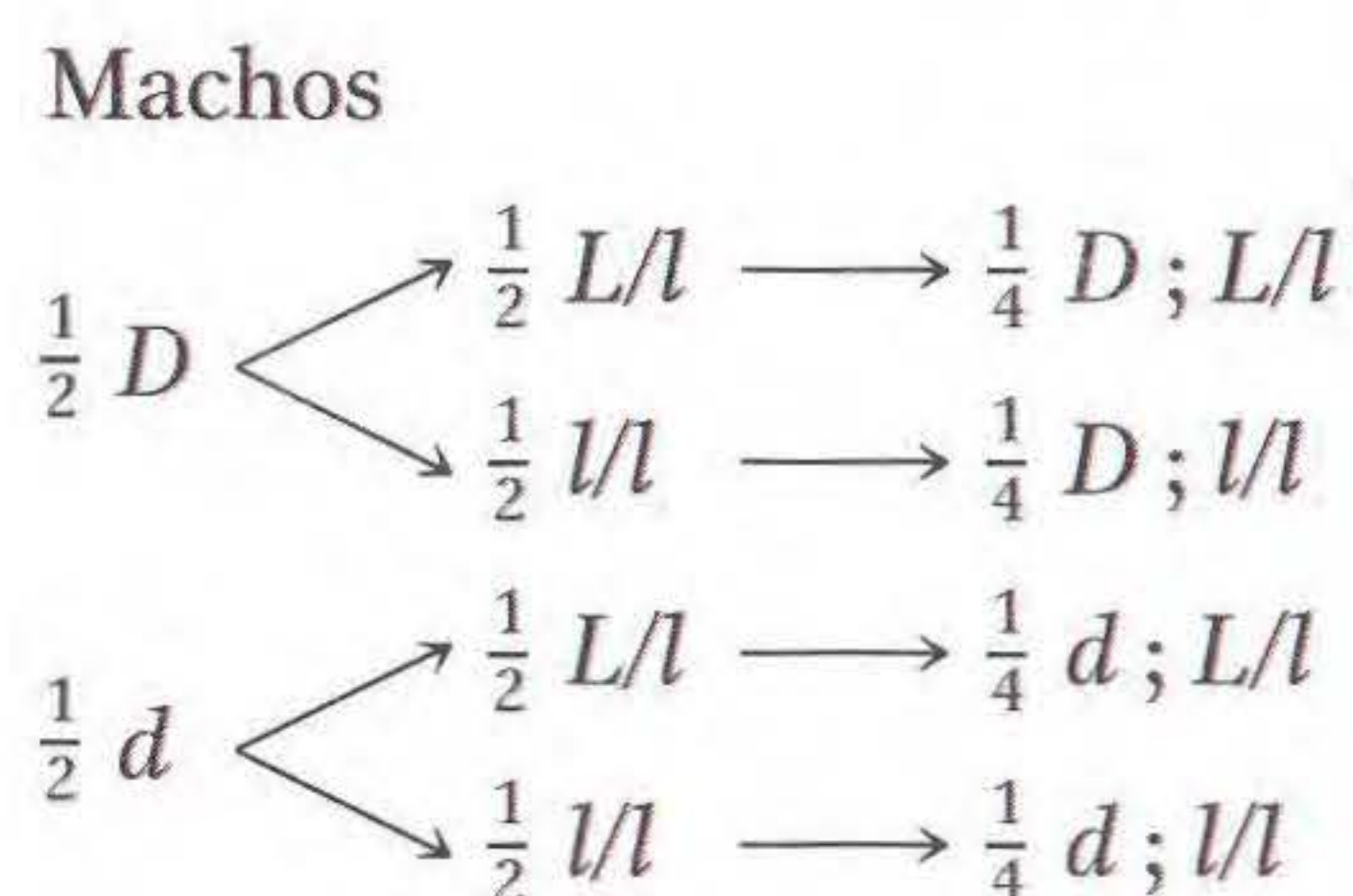
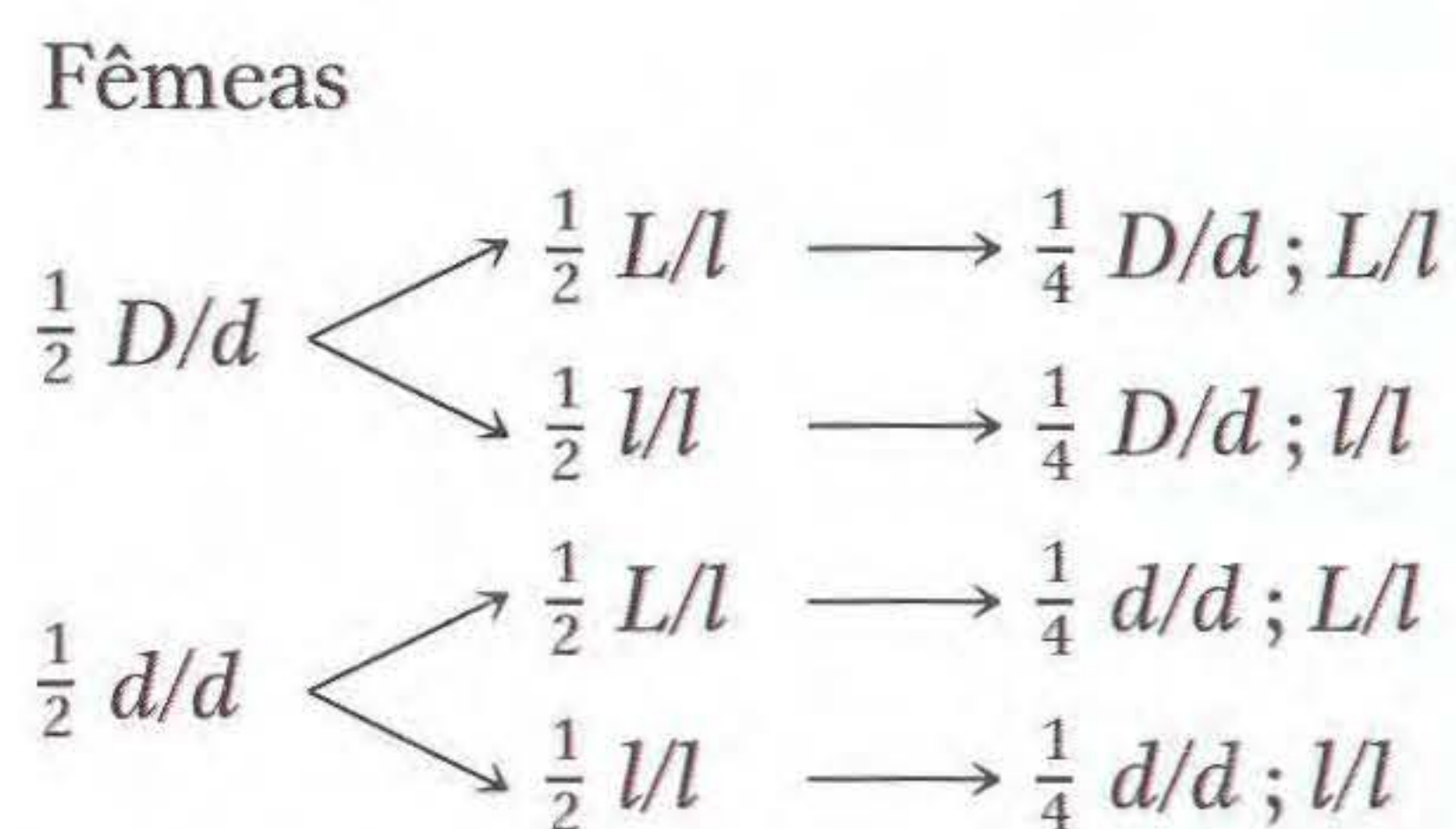
Assim, no geral, podemos representar o genitor feminino como $D/d ; L/l$ e o genitor masculino como $D ; L/l$. Então a prole seria



b. Geralmente, um bom modo de testar tal modelo é fazer um cruzamento e prever o resultado. Mas que cruzamento? Temos que prever algum tipo de proporção na prole, e assim é importante fazer um cruzamento a partir do qual uma única proporção fenotípica possa ser esperada. Note que usar uma da prole fêmea como genitora não atenderia nossas necessidades: não podemos dizer, pela observação do fenótipo de qualquer uma dessas fêmeas, qual é o seu genótipo. Uma fêmea com asas transparentes poderia ser D/D ou D/d , e uma com asas longas poderia ser L/L ou L/l . Seria bom cruzar a fêmea parental do cruzamento original com um filho fosco, aparado, pois todos os genótipos de ambos são especificados sob o modelo que criamos. De acordo com nosso modelo, esse cruzamento é:

$$D/d ; L/l \times d ; l/l$$

Desse cruzamento, prevemos



Problema resolvido 2. Considere três ervilhas lisas amarelas, marcadas como A, B e C. Cada uma foi cultivada em uma planta e cruzada com uma planta cultivada de uma ervilha verde rugosa. Exatamente 100 ervilhas produzidas de cada cruzamento foram distribuídas nas seguintes classes fenotípicas:

- A: 51 amarelas, lisas
49 verdes, lisas
- B: 100 amarelas, lisas
- C: 24 amarelas, lisas
26 amarelas, rugosas
25 verdes, lisas
25 verdes, rugosas

Quais eram de genótipos A, B e C? (Use símbolos gênicos de sua própria escolha; esteja seguro de definir cada um.)

SOLUÇÃO

Note que cada um dos cruzamentos é

amarela, lisa \times verde, rugosa
 \downarrow
prole

Como A, B e C foram todas cruzadas com a mesma planta, todas as diferenças entre as populações da prole devem ser atribuíveis a diferenças nos genótipos subjacentes de A, B e C.

Você deve lembrar muito dessas análises do capítulo, o que é bom, mas vejamos o quanto podemos deduzir dos dados. E quanto à dominância? O cruzamento-chave para deduzir a dominância é B. Aqui, o padrão de herança é

amarela, lisa \times verde, rugosa
 \downarrow
todas amarelas, lisas

Logo amarelo e liso devem ser fenótipos dominantes porque a dominância é literalmente definida pelo fenótipo de um híbrido. Agora sabemos que o genitor verde, liso, usado em cada cruzamento deve ser totalmente recessivo. Temos uma situação muito conveniente porque isso significa que cada cruzamento é um cruzamento-teste, que geralmente é o tipo mais informativo de cruzamento.

Voltando à prole de A, vemos uma proporção de 1:1 de amarela para verde. Essa proporção é uma demonstração da

Primeira Lei de Mendel (segregação igual) e mostra que, para a característica de cor, o cruzamento deve ter sido entre heterozigoto \times homozigoto recessivo. Se Y representar amarelo e y representar verde, temos

$$\begin{array}{c} Y/y \times y/y \\ \downarrow \\ \frac{1}{2} Y/y \text{ (amarelo)} \\ \frac{1}{2} y/y \text{ (verde)} \end{array}$$

Para a característica da forma, como toda a prole é lisa, o cruzamento deve ter sido de homozigota dominante \times homozigota recessiva. Se R representar lisa e r representar rugosa, temos

$$\begin{array}{c} R/R \times r/r \\ \downarrow \\ R/r \text{ (lisa)} \end{array}$$

Combinando as duas características, temos

$$\begin{array}{c} R/R ; Y/y \times y/y ; r/r \\ \downarrow \\ \frac{1}{2} Y/y ; R/r \\ \frac{1}{2} y/y ; R/r \end{array}$$

Agora, o cruzamento B ficou claro e deve ter sido

$$\begin{array}{c} Y/Y ; R/R \times y/y ; r/r \\ \downarrow \\ Y/y ; R/r \end{array}$$

porque qualquer heterozigose na ervilha B teria dado origem a vários fenótipos da prole, e não apenas um.

E quanto a C? Aqui, temos uma proporção de 50 amarelos : 50 verdes (1:1) e uma proporção de 49 lisas : 51 rugosas (também 1:1). Logo, ambos os genes na ervilha C devem ser heterozigotos, e o cruzamento C foi

$$Y/y ; R/r \times y/y ; r/r$$

que é uma boa demonstração da Segunda Lei de Mendel (distribuição independente de genes diferentes).

Como um geneticista analisou esses cruzamentos? Basicamente, do mesmo modo que fez com menos etapas intercalares. Possivelmente algo como isto: “amarelo e lisa dominante; segregação monogênica em A; B homozigotos dominantes; segregação independente de dois genes em C”.

Problemas

PROBLEMAS BÁSICOS

- Suponha distribuição independente e comece com uma planta que é diíbrida $A/a ; B/b$:
 - Que proporção fenotípica é produzida por sua autofecundação?
 - Que proporção genotípica é produzida por sua autofecundação?
 - Que proporção fenotípica é produzida por seu cruzamento-teste?

- Que proporção genotípica é produzida por seu cruzamento-teste?
- A mitose normal ocorre em uma célula diplóide de genótipo $A/a ; B/b$. Qual dos seguintes genótipos pode representar possíveis células-filhas?
 - $A ; B$
 - $a ; b$
 - $A ; b$
 - $a ; B$

- e. $A/A ; B/B$
 - f. $A/a ; B/b$
 - g. $a/a ; b/b$
3. Em um organismo diplóide de $2n = 10$, suponha que você pode marcar todos os centrômeros derivados do genitor feminino e todos os centrômeros derivados do genitor masculino. Quando esse organismo produz gametas, quantas combinações de centrômeros marcados femininos e masculinos são possíveis nos gametas?
 4. Em milho, o DNA em vários núcleos foi medido com base em sua absorção de luz. As medidas foram
0,7; 1,4; 2,1; 2,8 e 4,2

Que células podem ser usadas para essas medidas? (**Nota:** Nas plantas, o endosperma é geralmente triploide, $3n$.)

5. Desenhe uma mitose haplóide com o genótipo $a^+ ; b$.
6. Em musgo, os genes A e B são expressos apenas no gametófito. Um esporófito $A/a ; B/b$ produz gametófitos.
 - a. Que proporção de gametófitos será $A ; B$?
 - b. Se a fertilização é aleatória, que proporção de esporófitos na geração seguinte será $A/a ; B/b$?
7. Quando uma célula com o genótipo $A/a ; B/b ; C/c$ tendo todos os genes em pares separados de cromossomos se divide mitoticamente, quais são os genótipos das células-filhas?
8. Na levedura haplóide *Saccharomyces cerevisiae*, os dois tipos reprodutivos são conhecidos como MAT α e MAT a . Você cruza uma linhagem púrpura (ad^-) do tipo reprodutivo a com uma linhagem branca (ad^+) do tipo reprodutivo α . Se ad^- e ad^+ são alelos de um gene, e a e α são alelos de um gene herdado independentemente em um par de cromossomos separado, que prole você esperaria obter? Em que proporções?
9. Em camundongos, o nanismo é causado por um alelo recessivo ligado ao X, e a pelagem rosa é causada por um alelo autossômico dominante (pelagens são normalmente marrons). Se uma fêmea anã de uma linhagem pura é cruzada com um macho rosa de uma linhagem pura, quais serão as proporções fenotípicas na F_1 e na F_2 em cada sexo? (Invente e defina seus próprios símbolos gênicos.)
10. Suponha que você descobriu duas anomalias citológicas raras interessantes no cariótipo de um homem. (Um cariótipo é o complemento cromossômico total visível.) Existe um pedaço extra (satélite) em um dos cromossomos do par 4, e há um padrão anormal de corolação em um dos cromossomos do par 7. Supondo que todos os gametas desse homem sejam igualmente viáveis, que proporção de seus filhos terá o mesmo cariótipo que ele?
11. Suponha que a meiose ocorra no estágio transitório diplóide do ciclo de um organismo haplóide com número n de cromossomos. Qual a probabilidade de uma célula haplóide individual resultante da divisão meiótica ter um conjunto parental completo de centrômeros (isto é, um conjunto todo de um genitor ou todo do outro genitor)?
12. Imagine que o ano é 1868. Você é um habilidoso produtor de lentes trabalhando em Viena. Com suas superiores

novas lentes, você construiu um microscópio que tem resolução melhor do que qualquer outro disponível. No seu teste desse microscópio, você observa as células nos testículos de gafanhotos e fica fascinado com o comportamento de estranhas estruturas alongadas que você viu dentro de células em divisão. Um dia, na biblioteca, você lê uma publicação recente de G. Mendel sobre “fatores hipotéticos que ele diz explicar os resultados de alguns cruzamentos em ervilhas. Você é então surpreendido com a revelação dos paralelismos entre seus estudos com gafanhotos e os estudos de ervilha de Mendel, e resolve escrever-lhe uma carta. O que você escreveria? (Baseado em um idêia de Ernest Kroeker.)

13. De um suposto cruzamento-teste $A/a \times a/a$, no qual A representa vermelho e a branco, use o teste do χ^2 para descobrir qual dos possíveis resultados se ajustaria às expectativas:
 - a. 120 vermelhas, 100 brancas
 - b. 5.000 vermelhas, 5.400 brancas
 - c. 500 vermelhas, 540 brancas
 - d. 50 vermelhas, 54 brancas
14. Veja o quadrado de Punnett na Figura 3.4.
 - a. Quantos genótipos existem nos 16 quadrados da grade?
 - b. Qual a proporção genotípica subjacente a proporção fenotípica 9:3:3:1?
 - c. Você pode imaginar uma fórmula simples para o cálculo do número de genótipos da prole em cruzamentos diíbridos, triíbridos etc.? Repita para os fenótipos.
 - d. Mendel previu que, dentro de todas as classes fenotípicas menos uma no quadrado de Punnett, deve haver vários genótipos diferentes. Em particular, ele fez muitos cruzamentos para identificar os genótipos subjacentes do fenótipo lisa, amarela. Mostre dois modos diferentes que podem ser usados para identificar os vários genótipos subjacentes ao fenótipo lisa, amarela. (Lembre-se de que todas as ervilhas lisas, amarelas parecem idênticas.)
15. Supondo distribuição independente de todos os genes, desenvolva fórmulas que mostrem o número de classes fenotípicas e o número de classes genotípicas de autofecundação de uma planta heterozigota para n pares de genes.
16. **Nota:** A primeira parte deste problema foi introduzida no Capítulo 2. O raciocínio lógico é ampliado aqui.

Na planta *Arabidopsis thaliana*, um geneticista está interessado no desenvolvimento de tricomas (pequenas projeções) das folhas. Uma grande triagem revela que duas plantas mutantes (A e B) não têm tricomas, e esses mutantes parecem ser potencialmente úteis no estudo do desenvolvimento de tricomas. (Se elas foram determinadas por mutações monogênicas, então descobrir a função normal e anormal desses genes será instrutivo.) Cada planta ele cruzou com o tipo selvagem; em ambos os casos, a geração seguinte (F_1) teve tricomas normais. Ao serem autofecundadas plantas F_1 , as F_2 resultantes foram as seguintes:

F_2 do mutante A: 602 normais; 198 sem tricomas
 F_2 do mutante B: 267 normais; 93 sem tricomas

- a. O que esses resultados mostram? Inclua genótipos propostos de todas as plantas em sua resposta.
- b. Suponha que os genes estão situados em cromossomos separados. É produzida uma F_1 cruzando o mutante original A com o mutante original B. Essa F_1 é submetida a cruzamento-teste: Que proporção da prole do cruzamento-teste não terá tricomos?
17. Em cachorros, a cor de pelagem escura é dominante em relação a albino, e pêlo curto é dominante em relação a pêlo longo. Suponha que esses efeitos são causados por dois genes que se distribuem independentemente, e escreva os genótipos dos genitores em cada um dos cruzamentos mostrados aqui, nos quais D e A representam fenótipos escuro e albino, respectivamente, e S e L representam os fenótipos pelagem curta e pelagem longa.

Fenótipos parentais	Número de prole			
	D, S	D, L	A, S	A, L
a. D, S × D, S	89	31	29	11
b. D, S × D, L	18	19	0	0
c. D, S × A, S	20	0	21	0
d. A, S × A, S	0	0	28	9
e. D, L × D, L	0	32	0	10
f. D, S × D, S	46	16	0	0
g. D, S × D, L	30	31	9	11

Use os símbolos C e c para os alelos de cor de pelagem escura e albina e os símbolos S e s para os alelos de pelagem curta e longa, respectivamente. Suponha homozigose a menos que exista evidência em contrário. (O Problema 17 é reimpresso com permissão de Macmillan Publishing Co., Inc., de M. Strickberger, *Genetics*. Copyright 1968 by Monroe W. Strickberger.)

18. Em tomates, dois alelos de um gene determinam as características púrpura (P) versus verde (G) nos caules, e dois alelos de um gene separado independente determinam as características de folhas “cut” (C) versus “potato” (Po). Os resultados para cinco cruzamentos de fenótipos de plantas de tomate são os seguintes:

Reprodução	Fenótipos parentais	Número de prole			
		P, C	P, Po	G, C	G, Po
1	P, C × G, C	321	101	310	107
2	P, C × P, Po	219	207	64	71
3	P, C × G, C	722	231	0	0
4	P, C × G, Po	404	0	387	0
5	P, Po × G, C	70	91	86	77

- a. Determine que alelos são dominantes.
- b. Quais os genótipos mais prováveis para os genitores em cada cruzamento? (O Problema 18 é de A.M. Srb, R.D. Owen and R.S. Edgar, *General Genetics*, 2nd ed. Copyright 1965 by W.H. Freeman and Company.)
19. Um alelo mutante em camundongos causa uma cauda inclinada. Seis pares de camundongos foram cruzados.

Seus fenótipos e os da prole são mostrados no quadro seguinte. N é o fenótipo normal; B é o fenótipo inclinado. Deduza o modo de herança desse fenótipo.

Cruzamentos	Genitores		Prole	
	♀	♂	♀	♂
1	N	B	Todos B	Todos N
2	B	N	$\frac{1}{2}$ B, $\frac{1}{2}$ N	$\frac{1}{2}$ B, $\frac{1}{2}$ N
3	B	N	Todos B	Todos B
4	N	N	Todos N	Todos N
5	B	B	Todos B	Todos B
6	B	B	Todos B	$\frac{1}{2}$ B, $\frac{1}{2}$ N

- a. Ele é recessivo ou dominante?
- b. Ele é autossômico ou ligado ao sexo?
- c. Quais são os genótipos para todos os genitores e prole?
20. A cor de olho normal de *Drosophila* é vermelho, mas as linhagens nas quais todas as moscas têm olhos castanhos estão disponíveis. Similarmente, as asas são normalmente longas, mas existem linhagens com asas curtas. Uma fêmea de uma linhagem pura com olhos castanhos e asas curtas é cruzada com um macho de uma linhagem pura. A F_1 consiste em fêmeas normais e asas curtas. Uma F_2 é então produzida entrecruzando a F_1 . Ambos os sexos das moscas de F_2 mostram os seguintes fenótipos:

$\frac{3}{8}$ olhos vermelhos, asas longas
 $\frac{3}{8}$ olhos vermelhos, asas curtas
 $\frac{1}{8}$ olhos castanhos, asas longas
 $\frac{1}{8}$ olhos castanhos, asas curtas

Deduza a herança desses fenótipos; use símbolos genéticos claramente definidos de sua invenção. Diga os genótipos de todas as três gerações e as proporções genotípicas da F_1 e F_2 .



Solucionando o Problema 20

Antes de tentar uma solução para este problema, tente responder as seguintes perguntas:

1. O que a palavra “normal” significa neste problema?
2. O que a palavra “linhagem” significa?
3. Faça um esquema simples das duas linhagens parentais mostrando os olhos, asas e diferenças sexuais.
4. Quantas características diferentes existem neste problema?
5. Quantos fenótipos diferentes existem neste problema e que fenótipos vão com quais características?
6. Qual é o fenótipo completo das fêmeas F_2 chamadas de “normais”?
7. Qual o fenótipo completo dos machos F_1 chamados de asas “curtas”?
8. Cite as proporções fenotípicas F_2 para cada característica que você viu na resposta da questão 4.
9. O que as proporções fenotípicas F_2 lhe dizem?
10. Que padrão importante de herança distingue a herança ligada ao sexo da herança autossômica?

11. Os dados de F_2 mostram tal critério de distinção?
12. Os dados de F_1 mostram tal critério de distinção?
13. O que você pode aprender sobre dominância na F_1 ? e na F_2 ?
14. Que regras sobre o simbolismo de tipo selvagem você pode usar para decidir que símbolos alélicos inventar para esses cruzamentos?
15. O que significa “deduza a herança desses fenótipos”?

Agora tente resolver o problema. Se você for incapaz de fazê-lo, faça uma lista de perguntas sobre coisas que você não entendeu. Examine conceitos-chave no começo do capítulo e indague quais são relevantes para suas perguntas. Se esse enfoque não funcionar, examine as mensagens deste capítulo e pense sobre quais podem ser relevantes para suas perguntas.

21. Em uma população natural de plantas anuais, uma única planta é encontrada com aspecto pálido e tem folhas amareladas. A planta é colhida e levada para o laboratório. As taxas de fotossíntese são muito baixas. O pólen de uma planta normal escura e folhas verdes é usada para fertilizar flores emasculadas de uma planta amarelada. Resulta uma centena de sementes, das quais apenas 60 germinam. Todas as plantas resultantes são de aspecto amarelo-pálido.

- a. Proponha uma explicação genética para o padrão de herança.
- b. Sugira um teste simples para seu modelo.
- c. Explique a reduzida fotossíntese, palidez e aspecto amarelado.

22. Qual a base para a variação de cor verde e branca nas folhas de *Mirabilis*? Se for feito o cruzamento seguinte

$$\text{♀ variegada} \times \text{♂ verde}$$

que tipos de prole podem ser previstos? E quanto ao cruzamento recíproco?

23. Em *Neurospora*, o mutante *stp* exibe um crescimento errático pára-e-anda. O sítio mutante é conhecido como estando no mtDNA. Se uma linhagem *stp* é usada como genitor feminino em um cruzamento com uma linhagem normal agindo como macho, que tipo de prole pode ser esperado? E quanto à prole do cruzamento recíproco?

24. São estudadas duas plantas de milho. Uma é resistente (R) e a outra é suscetível (S) a certo fungo patogênico. São feitos os seguintes cruzamentos, com os resultados mostrados:

$$\begin{array}{lll} \text{♀ S} \times \text{♂ R} & \rightarrow & \text{toda a prole S} \\ \text{♀ R} \times \text{♂ S} & \rightarrow & \text{toda a prole R} \end{array}$$

O que você pode concluir sobre a localização dos determinantes genéticos de R e S?

25. Um suposto diíbrido em *Drosophila*, $B/b ; F/f$ é submetido a um cruzamento-teste com $b/b ; f/f$. (B = corpo preto; b = corpo marrom; F = cerdas bifurcadas; f = cerdas não-bifurcadas.) Os resultados são

Preto, bifurcado	230
Preto, não-bifurcado	210
Marrom, bifurcado	240
Marrom, não-bifurcado	250

Use o teste do χ^2 para determinar se esses resultados se ajustam aos resultados esperados do cruzamento-teste do híbrido hipotético.

26. Os seguintes números de prole são consistentes com os resultados esperados da autofecundação de uma planta suposta como sendo diíbrida de dois genes que se distribuem independentemente, $H/h ; R/r$? (H = folhas pilosas; h = folhas lisas; R = ovário redondo; r = ovário alongado.) Explique sua resposta.

pilosa, redondo	178
pilosa, alongado	62
lisa, redondo	56
lisa, alongado	24

27. Uma fêmea de mariposa escura é cruzada com um macho escuro. Todos os machos da prole são escuros, mas metade da prole de fêmeas são claras e as restantes são escuras. Proponha uma explicação para esse padrão de herança.

28. Em *Neurospora*, uma linhagem mutante chamada stopper (*stp*) surgiu espontaneamente. Stopper apresenta crescimento errático “pára e começa”, comparado com o crescimento ininterrupto de linhagens tipo selvagem. Em cruzamentos, foram encontrados os seguintes resultados:

$$\begin{array}{l} \text{♀ stopper} \times \text{♂ tipo selvagem} \rightarrow \text{prole toda stopper} \\ \text{♀ tipo selvagem} \times \text{♂ stopper} \rightarrow \text{prole toda tipo selvagem} \end{array}$$

- a. O que esses resultados sugerem quanto à localização da mutação stopper no genoma?

- b. De acordo com seu modelo para a parte a, que prole e proporções são previstas em óctades do seguinte cruzamento, incluindo uma mutação *nic3* situada no cromossomo VI?

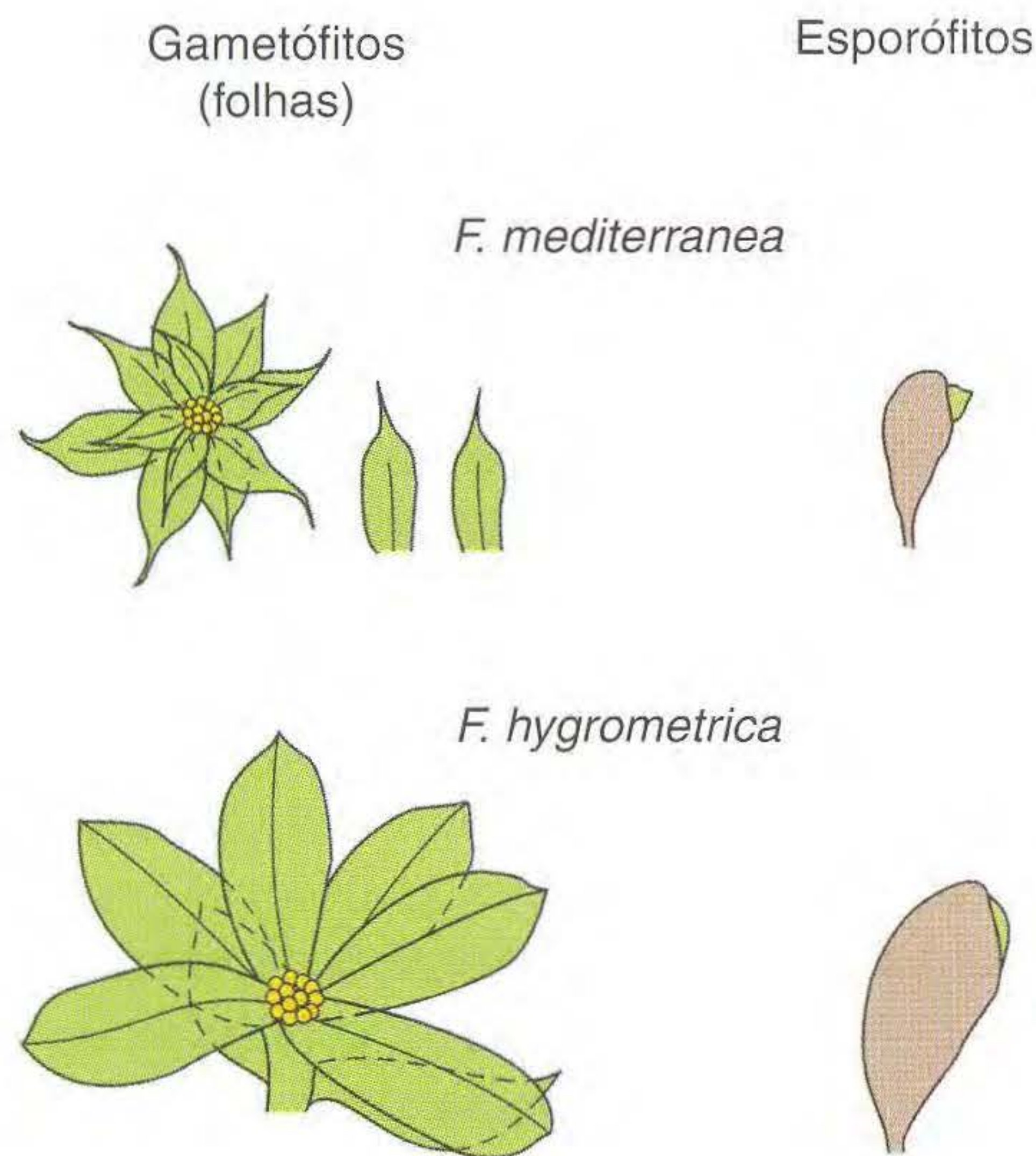
$$\text{♀ } stp \cdot nic3 \times \text{♂ tipo selvagem}$$

29. Em sistemas poligênicos, quantas classes fenotípicas correspondentes ao número de “doses” de poligenes são esperadas em autofecundações

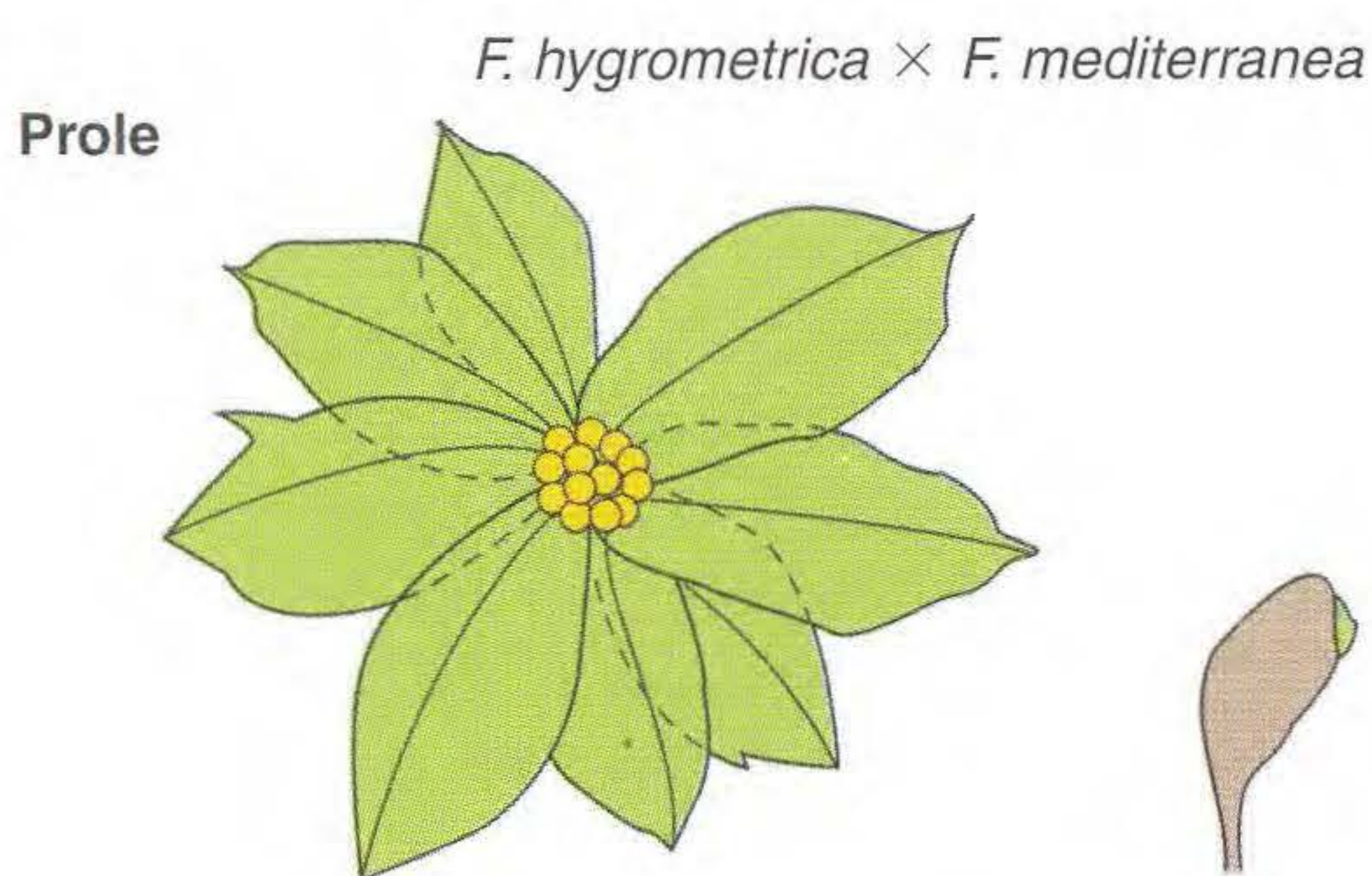
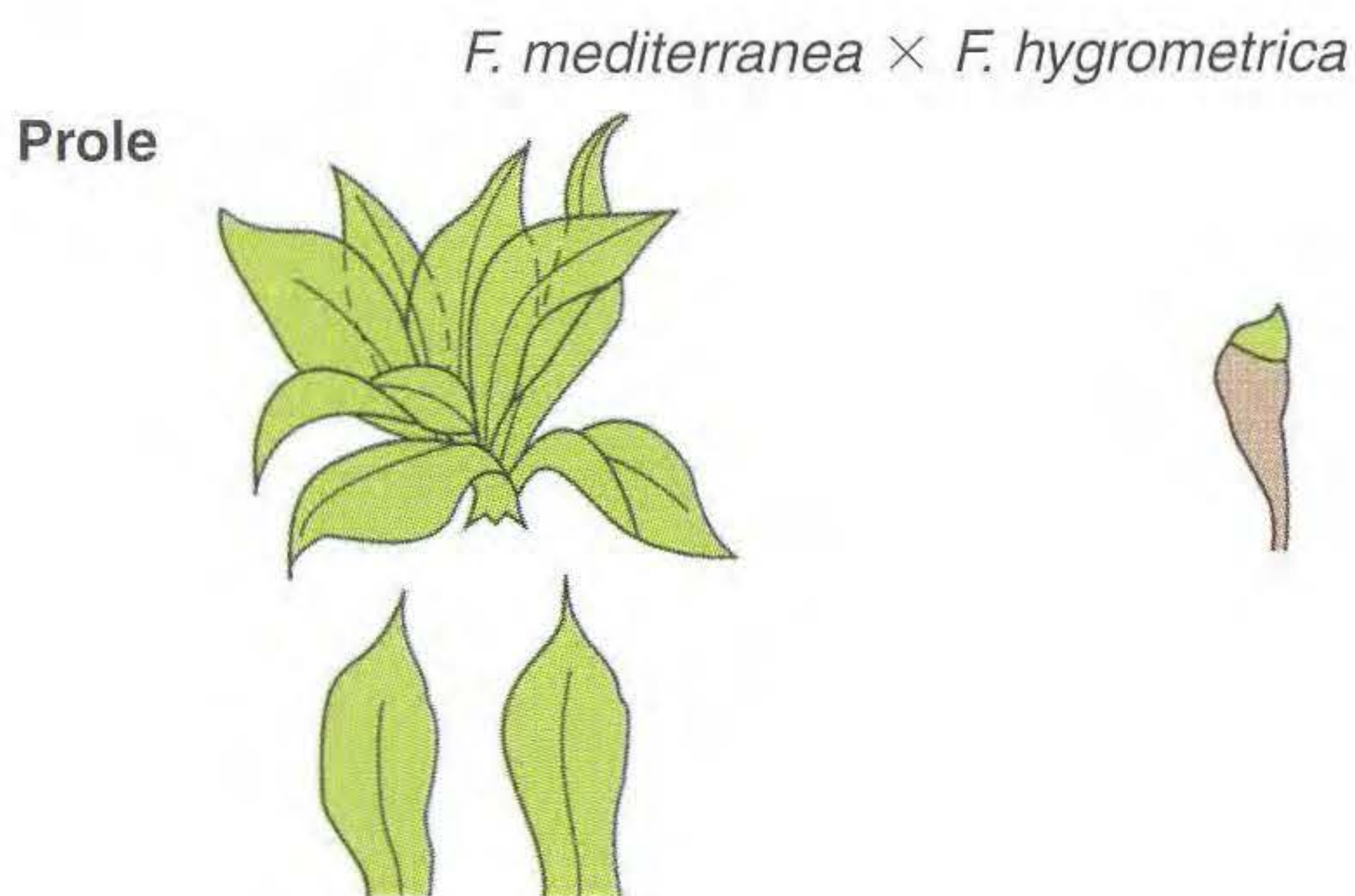
- a. de linhagens com quatro poligenes heterozigotos?
- b. de linhagens com seis poligenes heterozigotos?

30. Na autofecundação de um triíbrido poligênico $R_1/r_1 ; R_2/r_2 ; R_3/r_3$, use as regras do produto e da soma para calcular a proporção da prole apenas com uma “dose” de poligene.

31. Os cruzamentos recíprocos e autofecundações foram feitos entre duas espécies de musgos *Funaria mediterranea* e *F. hygrometrica*. Os esporófitos e as folhas dos gametófitos são mostradas no diagrama seguinte.



Os cruzamentos são escritos com o genitor feminino primeiro.

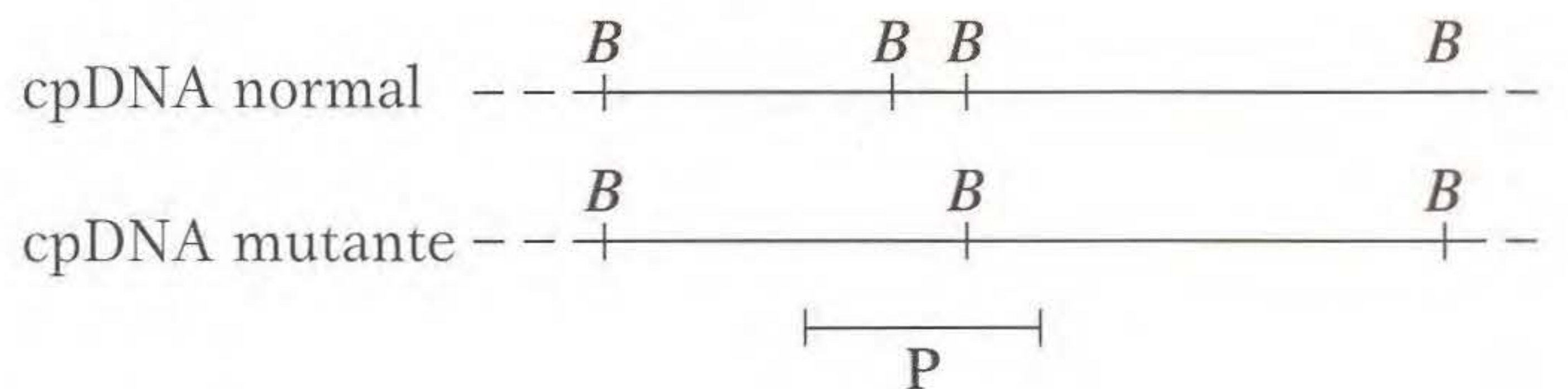


- Descreva os resultados apresentados, resumindo os principais achados.
- Proponha uma explicação dos resultados.
- Mostre como você testaria sua explicação; esteja certo de mostrar como ela pode ser diferenciada de outras explicações.

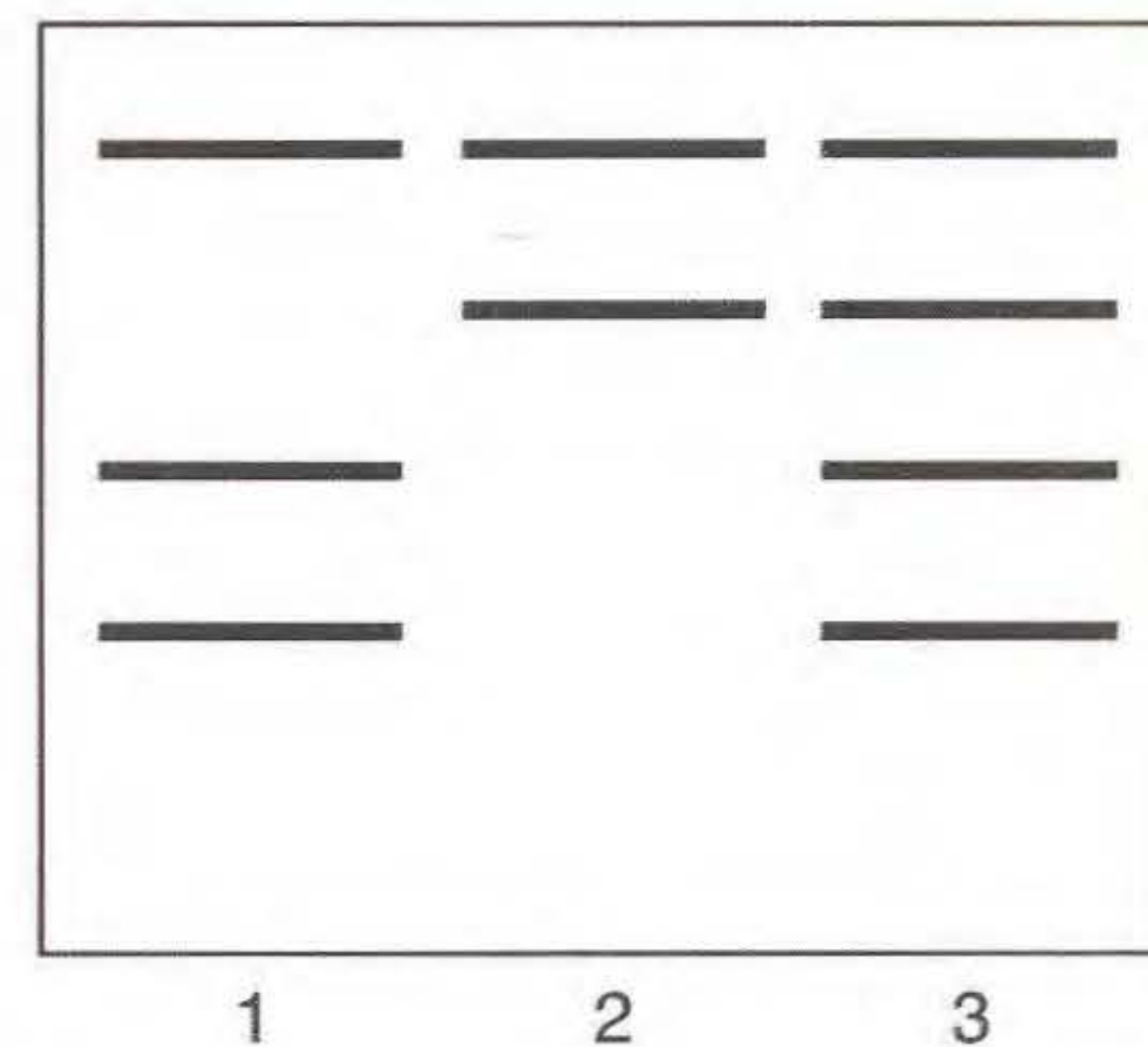
- Suponha que uma planta diplóide A tem um citoplasma geneticamente diferente do de uma planta B. Para estudar relações nucleocitoplasmáticas, você quer obter uma planta com o citoplasma da planta A e o genoma nuclear predominantemente da planta B. O que você faria para produzir tal planta?
- Você está estudando uma planta com tecido que tem setores tanto verdes quanto brancos. Você quer decidir

se esse fenômeno é devido (1) a uma mutação de cloroplasto do tipo considerado neste capítulo ou (2) a uma mutação nuclear dominante que inibe a produção de clorofila e está presente apenas em algumas camadas de tecidos da planta como um mosaico. Faça um esquema do enfoque experimental que você usaria para resolver este problema.

- No início do desenvolvimento de uma planta, uma mutação no cpDNA remove um sítio de restrição específico BgIII (B) como o seguinte:



Nessa espécie, o cpDNA é herdado maternamente. As sementes da planta são cultivadas, e a prole resultante é estudada quanto ao cpDNA. Os cpDNA são cortados com BgIII, e as transferências de Southern são hibridizadas com a sonda P mostrada. Os auto-radiogramas mostram três padrões de hibridização:



Explique a produção desses três tipos de sementes.

PROBLEMAS DESAFIADORES

- Você tem três jarras contendo bolinhas de gude, como segue:

jarra 1	600 vermelhas	e	400 brancas
jarra 2	900 azuis	e	100 brancas
jarra 3	10 verdes	e	990 brancas

- Se você, sem olhar, seleciona uma bola de cada jarro, calcule a probabilidade de obter

- uma vermelha, uma azul e uma verde.
- três brancas.
- uma vermelha, uma verde e uma branca.
- uma vermelha e duas brancas.
- uma colorida e duas brancas.
- pelo menos uma branca.

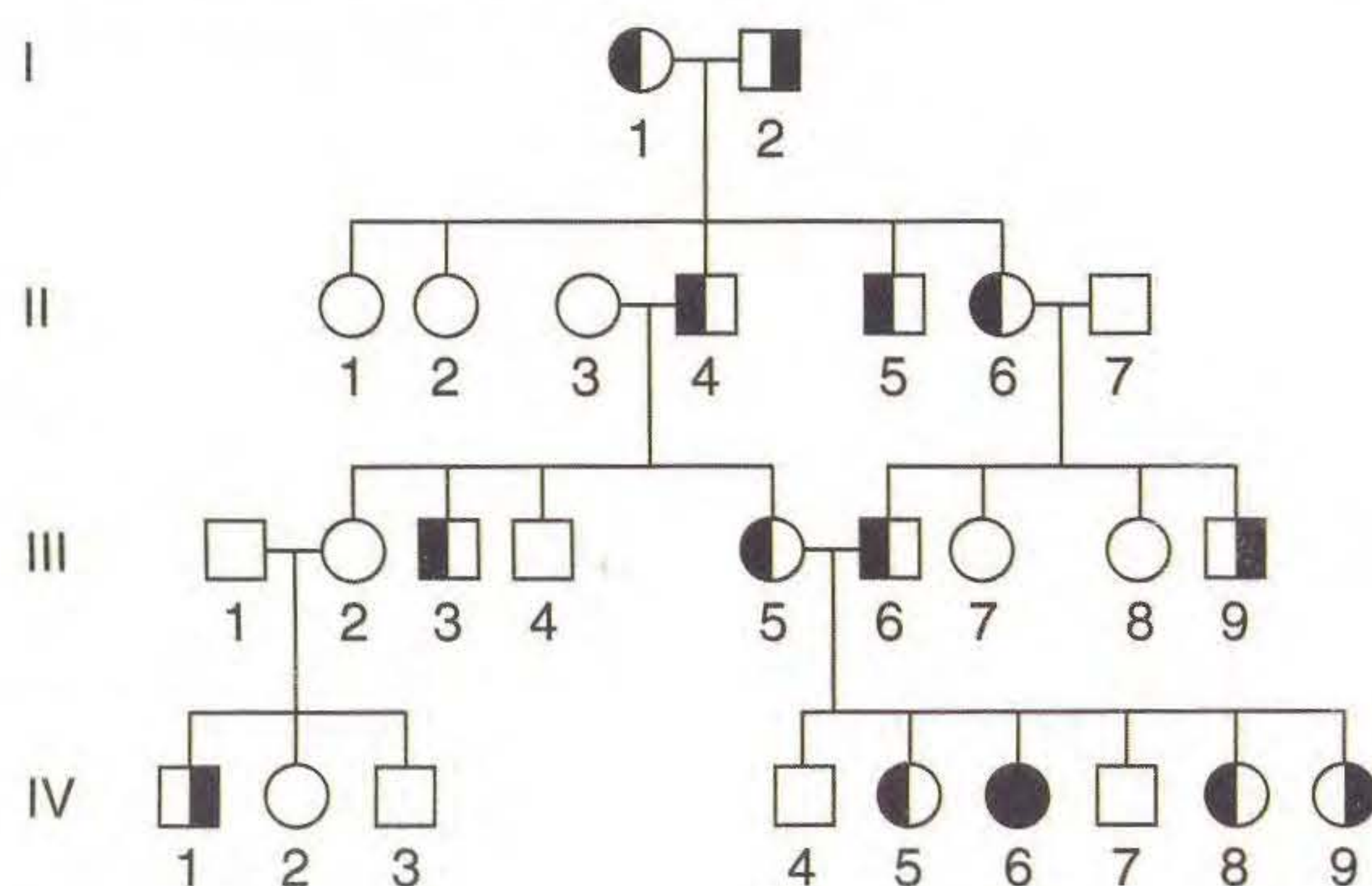
- Em uma certa planta, R = vermelha e r = branca. Você autofecunda um heterozigoto R/r a fim de obter uma planta branca para um experimento. Que número mínimo de sementes você deve cultivar para ter pelo menos 95% de certeza de obter pelo menos um indivíduo branco?

c. Quando uma mulher é injetada com um ovócito fertilizado *in vitro*, a probabilidade de esse implante ser bem-sucedido é de 20%. Se uma mulher é injetada com cinco zigotos simultaneamente, qual a probabilidade de ela ficar grávida? (Parte c é de Margaret Holm.)

36. Em tomates, o fruto vermelho é dominante em relação ao amarelo, o fruto biloculado é dominante sobre o multiloculado e a planta alta é dominante sobre a baixa. Um agricultor tem duas linhagens puras: (1) vermelho, biloculado, baixa e (2) amarelo, multiloculado, alta. Dessas duas linhagens, ele quer produzir uma nova linhagem pura para vender que é amarela, biloculada e alta. Como exatamente ele deve proceder? Mostre não só os cruzamentos a serem feitos, mas também quantas proles devem ser amostradas em cada caso.
37. Temos lidado, principalmente, apenas com dois genes, mas os mesmos princípios aplicam-se a mais de dois genes. Considere o seguinte cruzamento:

$$A/a ; B/b ; C/c ; D/d ; E/e \times a/a ; B/b ; c/c ; D/d ; e/e$$

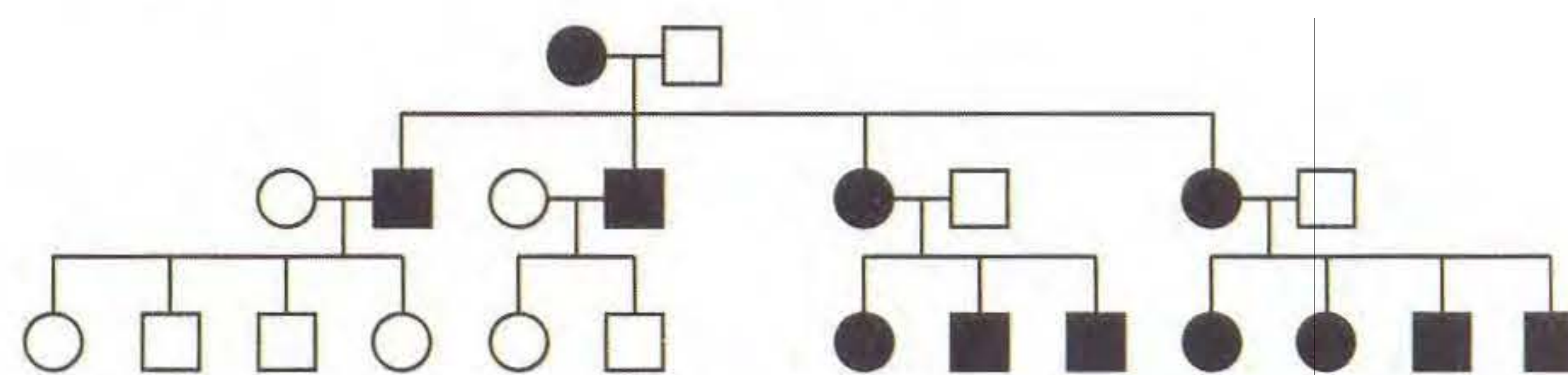
- a. Que proporção da prole *fenotipicamente* será parecida com (1) o primeiro genitor, (2) o segundo genitor, (3) ambos os genitores e (4) nenhum genitor?
- b. Que proporção da prole será *genotipicamente* igual ao (1) primeiro genitor, (2) segundo genitor, (3) ambos os genitores e (4) nenhum dos genitores?
- Suponha distribuição independente.
38. O heredograma acompanhante mostra o padrão de transmissão de dois fenótipos humanos raros: catarata e nanismo hipofisário. Os membros da família com catarata são mostrados com símbolos com a metade *esquerda* escura; os com nanismo hipofisário são indicados pela metade *direita* escura.



- a. Qual o modo de herança mais provável para cada um desses fenótipos? Explique.
- b. Cite os genótipos de todos os membros da geração III tanto quanto possível.
- c. Se uma reprodução hipotética ocorre entre IV-1 e IV-5, qual a probabilidade de o primeiro filho ser anão com catarata? Uma criança fenotipicamente normal? (O Problema 38 é de J. Kuspira and R. Bhambhani, *Compendium of Problems in Genetics*. Copyright 1994 by Wm. C. Brown.)
39. Um geneticista de milho tem três linhagens puras de genótipos $a/a ; B/B ; C/C$, $A/A ; b/b ; C/C$ e $A/A ; B/B ; c/c$.

Todos os fenótipos determinados por a , b e c aumentarão o valor de mercado do milho; logo, naturalmente, ele quer combiná-los em uma linhagem pura com o genótipo $a/a ; b/b ; c/c$.

- a. Programe um cruzamento efetivo que possa ser usado para obter a linhagem pura $a/a ; b/b ; c/c$.
- b. Em cada estágio, cite exatamente quais fenótipos serão selecionados e suas frequências esperadas.
- c. Há mais de um modo de obter o genótipo desejado? Qual o melhor modo?
- Suponha distribuição independente dos três pares de genes. (**Nota:** Milho pode se autofecundar ou polinizar facilmente.)
40. Em humanos, a visão em cores depende dos genes codificantes de três pigmentos. O gene para R (pigmento vermelho) e para G (pigmento verde) estão no cromossomo X, enquanto o gene B (pigmento azul) é autossômico. Uma mutação em qualquer um desses genes pode causar cegueira de cores. Suponha que um homem daltônico se case com uma mulher com visão em cores normal. Todos os seus filhos eram daltônicos, e todas as suas filhas, normais. Especifique os genótipos de ambos os genitores e todos os filhos possíveis, explicando seu raciocínio. (Desenhar um heredograma provavelmente será útil.) (Problema 40 é de Rosemary Redfield.)
41. Considere o heredograma acompanhante para uma doença muscular humana rara.



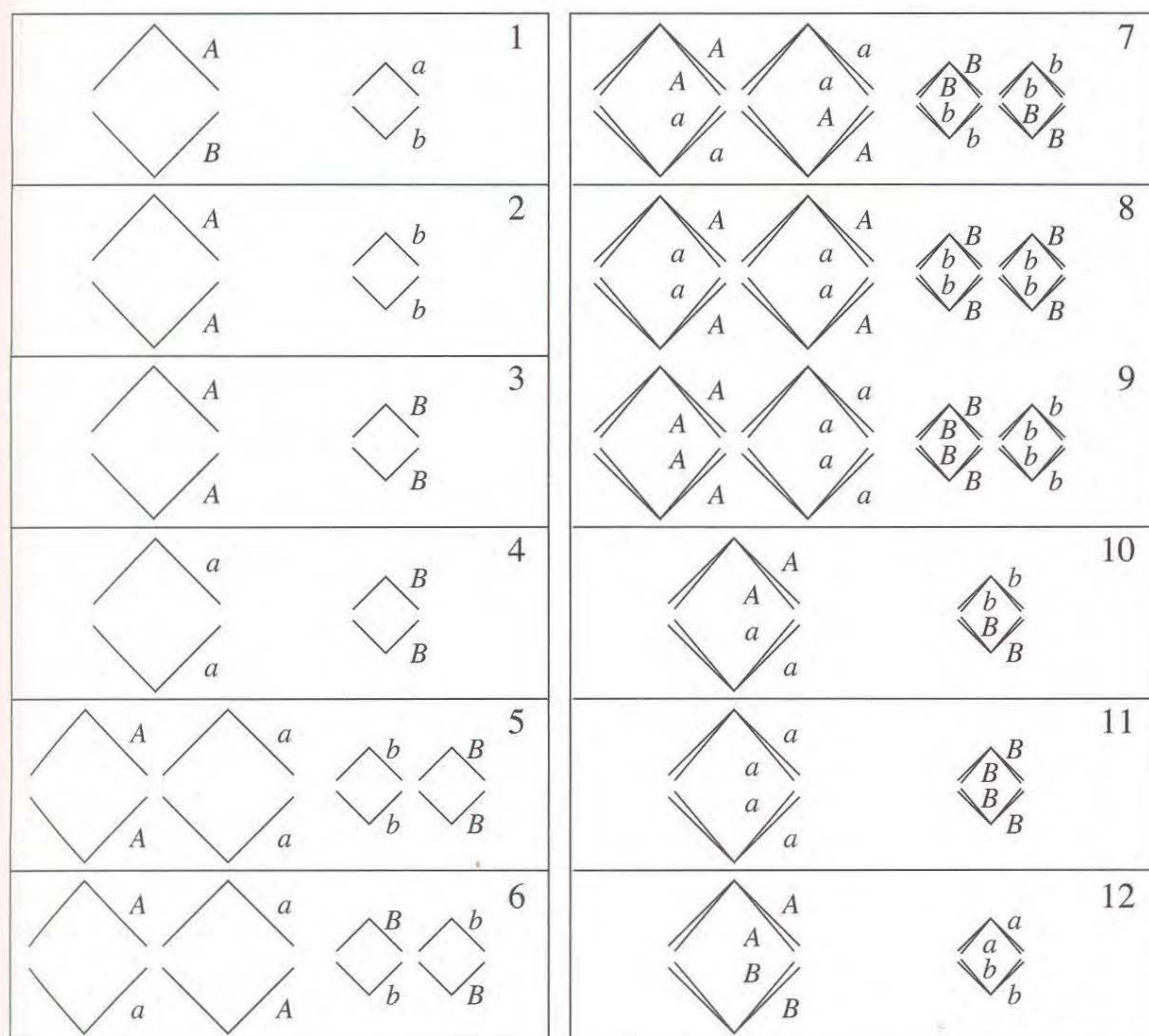
- a. Que característica incomum distingue esse heredograma dos estudados antes neste capítulo?
- b. Onde você acha que o DNA mutante responsável por esse fenótipo reside na célula?
42. A planta *Haplopappus gracilis* tem $2n$ de 4. Uma cultura de células diplóides foi estabelecida e, na fase S pré-mitótica, um nucleotídeo radioativo foi adicionado e incorporado ao DNA recém-sintetizado. As células foram então removidas da radioatividade, lavadas e deixadas para continuar a mitose. Os cromossomos radioativos, ou cromátides, podem ser detectados colocando-se uma emulsão fotográfica nas células; os cromossomos radioativos, ou cromátides, apresentam-se cobertos com pontos de prata da emulsão. (Os cromossomos "tiram sua própria fotografia".) Desenhe os cromossomos na prófase e telófase da primeira e segunda divisões mitóticas após o tratamento radioativo. Se eles são radioativos, mostre em seu diagrama. Se existem várias possibilidades, mostre-as também.
43. Na espécie do Problema 42, você pode introduzir radioatividade por injeção nas anteras na fase S antes da meiose. Desenhe os quatro produtos da meiose com seus cromossomos e mostre quais são radioativos.
44. As duplas hélices do DNA dos cromossomos podem ser particularmente deselcoicizadas *in situ* por tratamentos

especiais. Esse padrão de radioatividade é esperado se tal preparação for banhada em uma sonda radioativa para

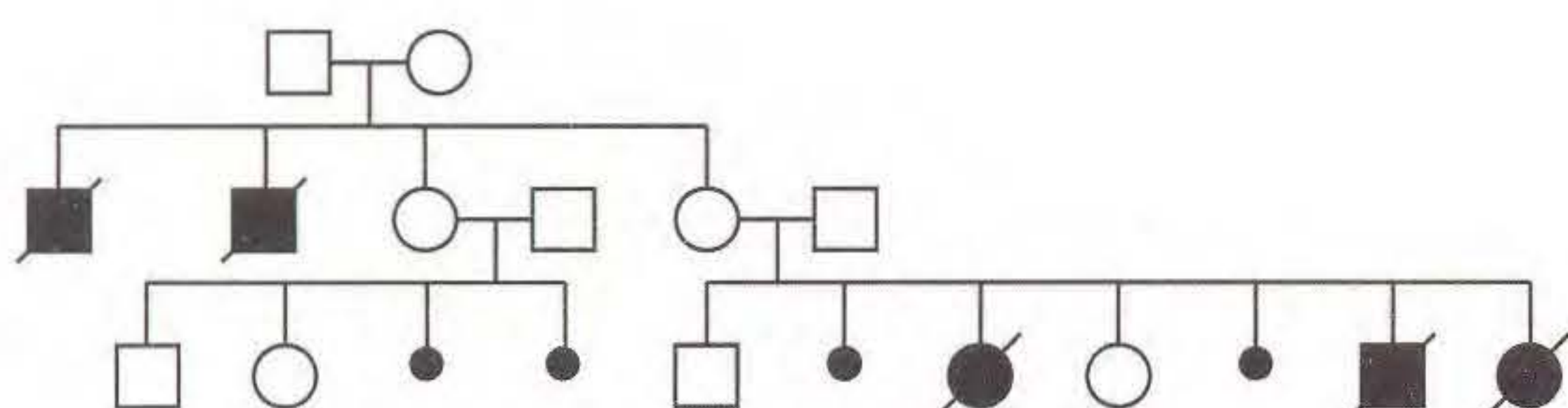
- um único gene?
- DNA repetitivo disperso?
- DNA ribossômico?
- DNA telomérico?
- DNA heterocromático com repetição única?

45. Se o DNA genômico é cortado com uma enzima de restrição e fracionado por tamanho pela eletroforese, que padrão de hibridização de Southern é esperado para as sondas citadas no Problema 44?

46. A planta *Haplopappus gracilis* é diplóide e $2n = 4$. Há um par de cromossomos longos e um par de curtos. Os diagramas acompanhantes (numerados de 1 a 12) representam anáfases (estágio de “separação”) de células individuais na meiose ou mitose em uma planta que é geneticamente um diíbrido ($A/a ; B/b$) para genes em cromossomos diferentes. As linhas representam cromossomos ou cromátides, e as pontas dos V representam centrômeros. Em cada caso, indique se o diagrama representa uma célula em meiose I, meiose II ou mitose. Se um diagrama mostra uma situação impossível, indique qual.



47. O heredograma a seguir mostra a recorrência de uma rara doença neurológica (símbolos grandes escuros) e os abortos fetais espontâneos (pequenos símbolos pretos) em uma família. (Uma barra significa que a pessoa já faleceu.) Cite uma explicação para o heredograma com relação à segregação citoplasmática de mitocôndrias defeituosas.



48. Um homem tem braquidactilia (dedos muito curtos; autossômica dominante rara) e sua esposa não. Ambos sentem o gosto da feniltiocarbamida (alelo comum; autossômico dominante), mas suas mães não.

a. Cite os genótipos do casal.

Se os genes segregam-se independentemente e o casal tem quatro filhos, qual a probabilidade de

- todos eles terem braquidactilia?
- nenhum ter braquidactilia?
- todos sentirem o gosto?
- nenhum deles sentir o gosto?
- todos eles terem braquidactilia e sentirem o gosto?
- nenhum deles ter braquidactilia nem sentir o gosto?
- pelo menos um ter braquidactilia e sentir o gosto?

49. Uma forma de esterilidade masculina em milho é transmitida maternamente. As plantas de uma linhagem com esterilidade masculina cruzada com pólen normal deu plantas com esterilidade masculina. Além disso, algumas linhagens de milho são conhecidas levando um alelo dominante nuclear restaurador (Rf) que recupera a fertilidade do pólen nas linhagens com esterilidade masculina.

a. Uma pesquisa mostrou que a introdução de alelos restauradores em linhagens com esterilidade masculina não altera nem afeta a manutenção dos fatores citoplasmáticos para esterilidade masculina. Que tipo de resultados de pesquisa levaria a tal conclusão?

b. Uma planta com esterilidade masculina é cruzada com pólen de uma planta homozigota para Rf . Qual o genótipo da F_1 ? O fenótipo?

c. As plantas de F_1 da parte b são usadas como fêmeas em um cruzamento-teste com pólen de uma planta normal (rf/rf). Quais são os resultados desse cruzamento-teste? Cite os genótipos e fenótipos e o tipo de citoplasma.

d. O alelo restaurador já descrito pode ser chamado de $Rf-1$. Outro restaurador dominante, $Rf-2$, foi encontrado. $Rf-1$ e $Rf-2$ estão situados em cromossomos diferentes. Um ou ambos os alelos restauradores darão fertilidade de pólen. Com o uso de uma planta com esterilidade do macho como testador, qual seria o resultado de um cruzamento no qual o genitor masculino é

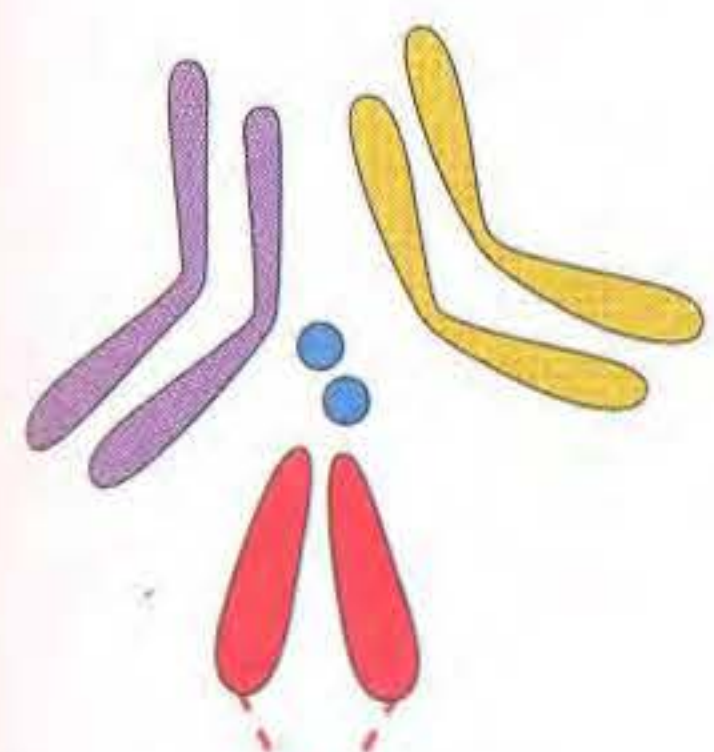
(i) heterozigoto em ambos os loci restauradores?

(ii) homozigoto dominante em um locus restaurador e homozigoto recessivo no outro?

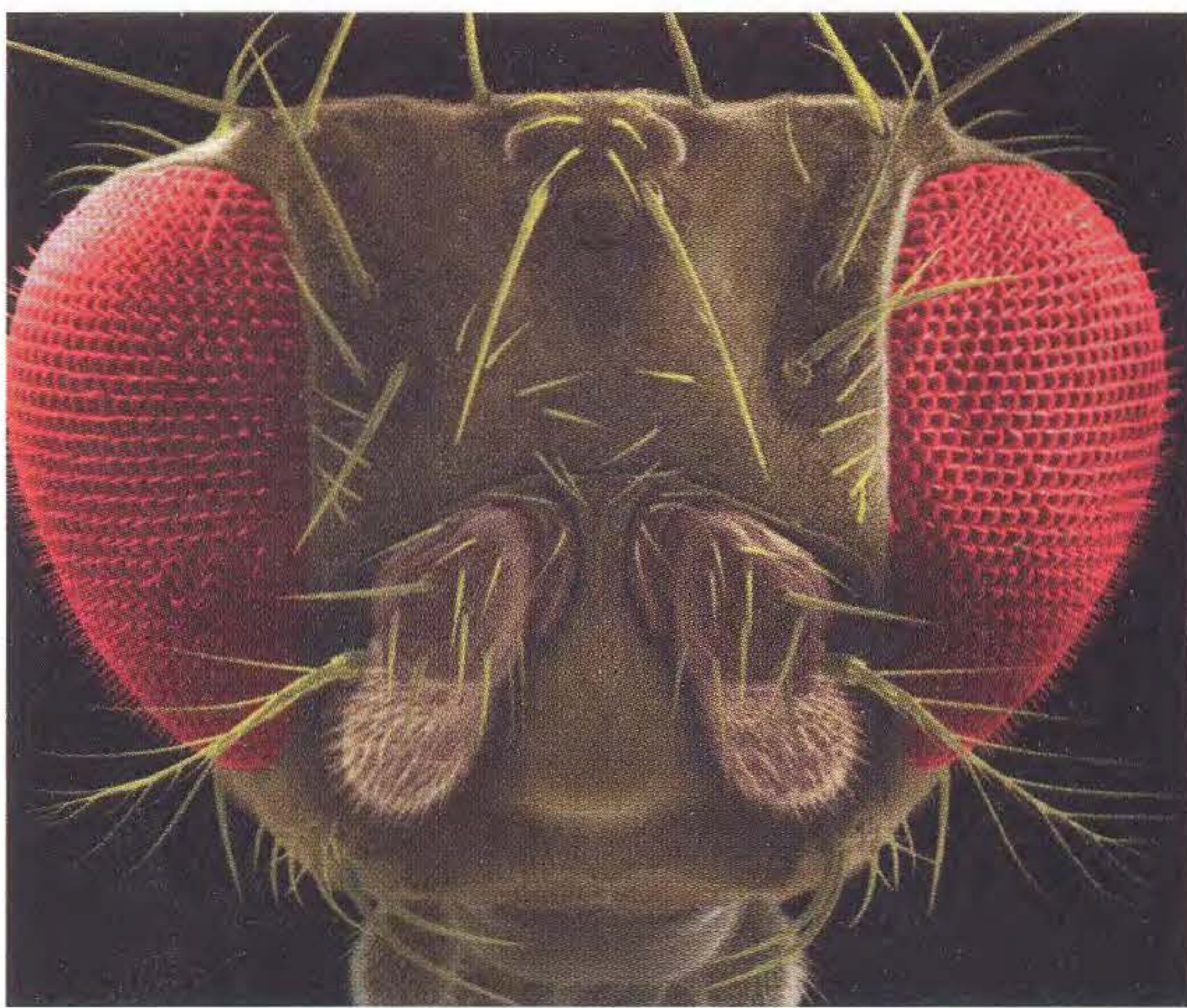
(iii) heterozigoto em um locus restaurador e homozigoto recessivo no outro?

(iv) heterozigoto em um locus restaurador e homozigoto dominante no outro?

Mapeamento de Cromossomos Eucarióticos por Recombinação



0.0	Yellow body
1.5	Scute bristles
3.0	White eyes
5.5	Facet eyes
7.5	Echinus eyes
13.7	Ruby eyes
20.0	Crossveinless wings
21.0	Cut wings
27.7	Singed bristles
33.0	Lozenge eyes
36.1	Vermillion eyes
43.0	Miniature wings
44.0	Sable body
56.7	Garnet eyes
57.0	Forked bristles
59.5	Bar eyes
62.5	Fused veins
66.0	Carnation eyes
	Bobbed hairs



À esquerda, um mapa baseado em recombinação de um dos cromossomos de *Drosophila* (o organismo na imagem à direita), mostrando os loci dos genes cujas mutações produzem fenótipos conhecidos. (Dennis Kunkel Microscopy, Inc.)

Algumas das perguntas que os geneticistas desejam responder sobre o genoma são: *Que genes* estão presentes no genoma? *Que funções* eles têm? *Que posições* eles ocupam nos cromossomos? Sua perseguição da terceira pergunta é de modo amplo chamada de mapeamento. O mapeamento é o principal foco deste capítulo, mas existem três questões inter-relacionadas, como veremos mais adiante no capítulo.

Todos temos uma sensação diária da importância dos mapas em geral, e, de fato, todos nós os usamos alguma vez em nossas vidas para encontrar um caminho. O foco relevante deste capítulo é que, em algumas situações, vários mapas precisam ser usados simultaneamente. Um bom exemplo na vida cotidiana é percorrer a densa rede de ruas e prédios em uma cidade tal como Londres, Inglaterra. Um mapa de ruas que mostra o esquema geral é uma necessidade. Entretanto, o mapa de ruas é usado por turistas e moradores de Londres em conjunto com outro mapa: o do sistema subterrâneo

Perguntas Fundamentais

- Que processos celulares produzem uma recombinação de genes ligados?
- Que frequências de recombinação são diagnósticas de ligação?
- Como pode a análise de frequências de recombinação gerar um mapa cromossômico?
- Como os mapas de recombinação são usados em conjunto com os mapas físicos de DNA?

Tópicos

- 4.1 Diagnóstico de ligação
- 4.2 Mapeamento por frequência de recombinação
- 4.3 Mapeamento com marcadores moleculares
- 4.4 Mapeamento de centrômero com tétrades lineares
- 4.5 Uso do teste do qui-quadrado para testar análise de ligação
- 4.6 Uso de valores lod para avaliar a ligação em heredogramas humanos
- 4.7 Cômputo de crossings múltiplos não-vistos
- 4.8 Uso de mapas baseados em recombinação em conjunto com mapas físicos

do metrô. O sistema subterrâneo é tão complexo e similar a uma macarronada que, em 1933, um engenheiro de circuito elétrico chamado Harry Beck desenhou um mapa (um pouco distorcido) que até hoje é um ícone de Londres. Os mapas de ruas e subterrâneos de Londres são comparados na Figura 4.1. Note que as posições das estações subterrâneas e as distâncias exatas entre elas não têm interesse em si mesmas, exceto como um meio de chegar ao destino de interesse, tal como a abadia de Westminster. Veremos paralelos aos mapas de Londres quando são usados mapas cromossômicos para indicar “destinos” determinados ou genes específicos. Primeiro, vários tipos diferentes de mapas cromossômicos em geral são necessários, e devem ser usados em conjunto; segundo, os mapas que contêm distorções ainda são úteis; e terceiro, muitos sítios em um mapa cromossômico são marcados apenas porque são úteis em tentar situar outros pontos que são os de real interesse.

A obtenção de um mapa das posições dos genes nos cromossomos é uma tarefa que ocupou milhares de geneticistas nos últimos 80 anos. Por que isso é tão importante? Existem vários motivos:

1. A posição do gene é uma informação crucial necessária para construir *genótipos complexos* necessários para fins experimentais ou aplicações comerciais. Por exemplo, no Capítulo 6, veremos casos nos quais combinações alélicas especiais devem ser juntadas para explorar a interação gênica.
2. Conhecer a posição ocupada por um gene fornece um modo de enfocar sua *estrutura e função*. A posição de um gene pode ser usada para defini-lo a nível de DNA. Por

sua vez, a sequência de DNA de um gene tipo selvagem ou seu alelo mutante é uma parte necessária para deduzir sua função subjacente.

3. Os genes presentes e sua disposição nos cromossomos são, em geral, ligeiramente diferentes em espécies correlatas. Por exemplo, o longo cromossomo humano número 2 é dividido em dois cromossomos mais curtos nos grandes macacos. Comparando tais diferenças, os geneticistas podem deduzir os *mecanismos genéticos evolutivos* pelos quais esses genomas divergiram. Assim, os mapas cromossômicos são úteis para interpretar os mecanismos da evolução.

A disposição dos genes nos cromossomos é representada diagramaticamente como um **mapa cromossômico** unidimensional, mostrando as posições dos genes, conhecidas como **loci** (singular locus), e as distâncias entre esses loci são baseadas em algum tipo de escala. Dois tipos básicos de mapas cromossômicos são atualmente usados em genética; eles são montados por procedimentos bem diferentes, embora sejam usados de modo complementar. Os *mapas baseados em recombinação*, que são o tópico deste capítulo, mapeiam os loci dos genes que foram identificados por fenótipos mutantes mostrando herança monogênica. Os *mapas físicos* (veja Capítulo 13) mostram os genes como segmentos dispostos ao longo da molécula de DNA que constitui um cromossomo. Esses mapas mostram visões diferentes do genoma, mas, como os mapas de Londres, eles podem ser usados juntos para chegar a uma compreensão de qual é a função de um gene a nível molecular e como essa função influencia o fenótipo.

Dois mapas são melhores do que um

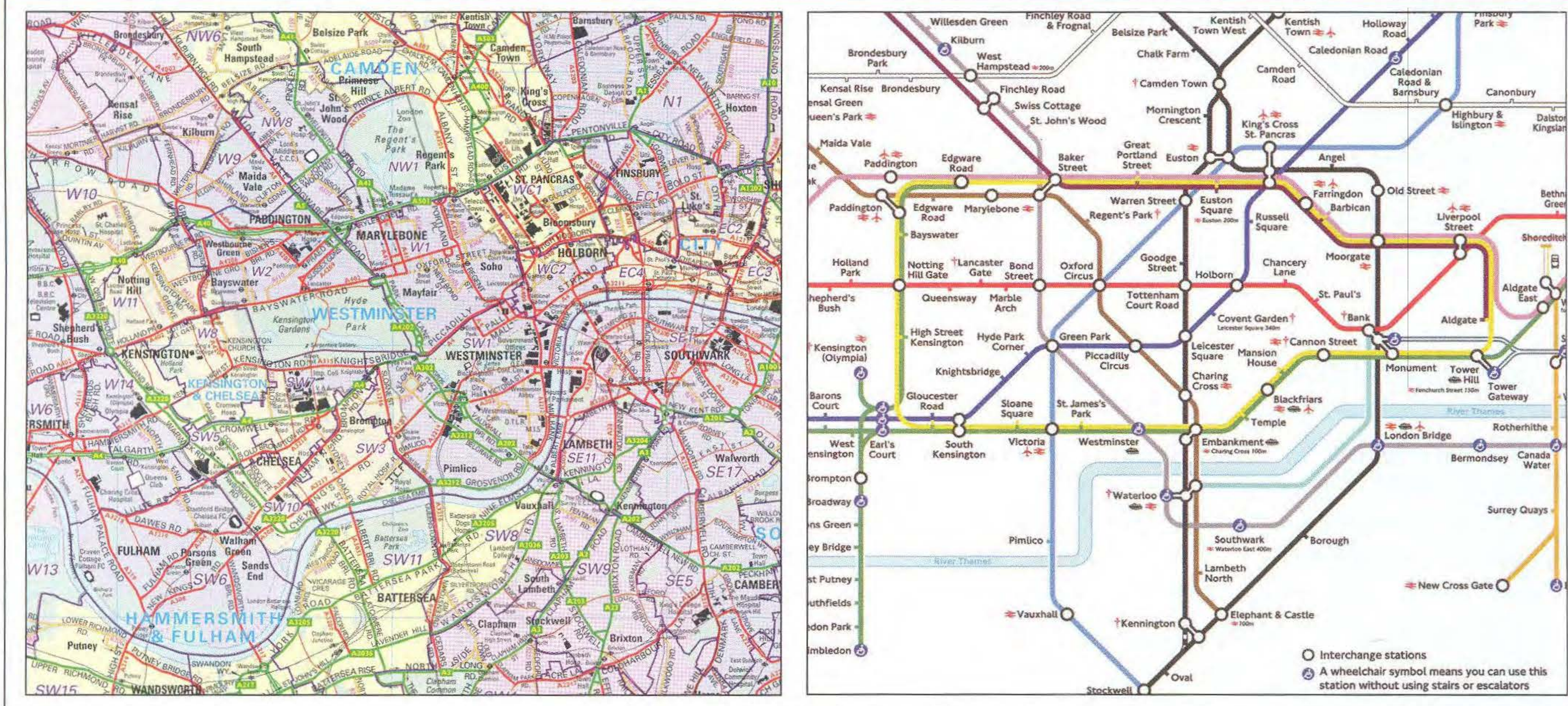


FIG. 4.1 Esses mapas de Londres ilustram o princípio de que, geralmente, são necessárias várias etapas para chegar ao destino de interesse. O mapa da Underground Railway (“the Tube”) é usado para chegar ao destino de interesse tal como o endereço de uma rua, mostrado no mapa de ruas. Em genética, dois tipos diferentes de mapas do genoma são geralmente úteis para localizar um gene, levando a uma compreensão de sua estrutura e função. [(esquerda) HarperCollins Publishers Ltd. (direita) © Transport for London. Reg. User nº 06/E/1632.]

Os mapas genéticos são úteis para a construção de uma linhagem, para interpretar os mecanismos evolutivos e para descobrir a função desconhecida de um gene. Descobrir a função de um gene é facilitada por integrar a informação baseada em recombinação e mapas físicos.

4.1 Diagnóstico de Ligação

Os **mapas de recombinação** dos cromossomos são geralmente montados com dois ou três genes de cada vez, com o uso de um método chamado análise de ligação. Quando os geneticistas dizem que dois genes são **ligados**, eles querem dizer que os loci desses genes estão no mesmo cromossomo, e, assim, os alelos de qualquer homólogo estão fisicamente unidos (ligados) pelo DNA entre eles. O modo pelo qual os primeiros geneticistas deduziram a ligação é um meio útil de introduzir a maioria das idéias principais e procedimentos na análise.

Uso de frequência de recombinantes para reconhecer ligação

No início dos anos de 1900, William Bateson e R. C. Punnett (de onde vem o nome quadrado de Punnett) estavam estudando a herança de dois genes em ervilhas-de-cheiro. Em uma autofecundação padrão de um diíbrido de F_1 , a F_2 não mostrou a proporção 9:3:3:1 prevista pelo princípio de distribuição independente. De fato, Bateson e Punnett notaram que certas combinações de alelos apresentavam-se com mais frequência do que o esperado, como se fossem fisicamente ligadas de algum modo. Entretanto, não tinham explicação para essa descoberta.

Mais tarde, Thomas Hunt Morgan encontrou um desvio similar da Segunda Lei de Mendel quando estudava dois genes autossômicos em *Drosophila*. Morgan propôs a ligação como uma hipótese para explicar o fenômeno de aparente associação de alelos.

Vejamos alguns dados de Morgan. Um dos genes afetava a cor de olho (pr , púrpura, e pr^+ , vermelho), e o outro gene afetava o tamanho da asa (vg , vestigial, e vg^+ , normal). Os alelos de ambos os genes são dominantes. Morgan fez um cruzamento para obter diíbridos e, então, acompanhar com um cruzamento-teste:

P $pr/pr \cdot vg/vg \times pr^+/pr^+ \cdot vg^+/vg^+$

Gametas $pr \cdot vg \quad pr^+ \cdot vg^+$

Diíbrido de F_1 $pr^+/pr \cdot vg^+/vg$

Cruzamento-teste:

♀ $pr^+/pr \cdot vg^+/vg$ × ♂ $pr/pr \cdot vg/vg$

Fêmea diíbrida de F_1 Macho testador

O uso de Morgan do cruzamento-teste é importante. Como o genitor testador contribui com gametas levando apenas alelos recessivos, os fenótipos da prole revelam diretamente os alelos contribuídos pelos gametas do genitor diíbrido, como

descrito nos Capítulos 2 e 3. Assim, a análise pode se concentrar na meiose em um genitor (o diíbrido) e, essencialmente, esquecer a meiose no outro (o testador). Em contraste, em um F_1 *autofecundado*, existem *dois* conjuntos de meiose a considerar na análise da prole: um no genitor masculino e um no feminino.

Os resultados do cruzamento-teste de Morgan foram os seguintes (citados como classes gaméticas do diíbrido):

$pr^+ \cdot vg^+$	1.339
$pr \cdot vg$	1.195
$pr^+ \cdot vg$	151
$pr \cdot vg^+$	154
	<hr/> 2.839

Obviamente, esses números desviam-se drasticamente da previsão mendeliana de uma proporção 1:1:1:1. As duas primeiras combinações de alelos estão, em grande maioria, indicando claramente que estão associadas ou “ligadas”.

Outro modo útil de avaliar o cruzamento-teste é considerando a percentagem de recombinantes na prole. Por definição, os recombinantes no presente cruzamento são de dois tipos $pr^+ \cdot vg$ e $pr \cdot vg^+$ porque, claramente, eles não são os dois genótipos contribuídos para o diíbrido de F_1 pelas moscas parentais homozigotas originais (mais exatamente, por seus gametas). Vemos que os dois tipos recombinantes são aproximadamente iguais em frequência (151 ~ 154). Seu total é 305, que é uma frequência de $(305/2.839) \times 100$, ou 10,7%. Esses dados podem fazer sentido, como Morgan fez, postulando que os genes estavam ligados no mesmo cromossomo, e assim as combinações alélicas parentais são mantidas juntas na maioria da prole. No diíbrido, a conformação alélica deve ter sido a seguinte:

$$\frac{pr^+ \quad vg^+}{pr \quad vg}$$

A tendência de genes ligados serem herdados em bloco é ilustrada na Figura 4.2.

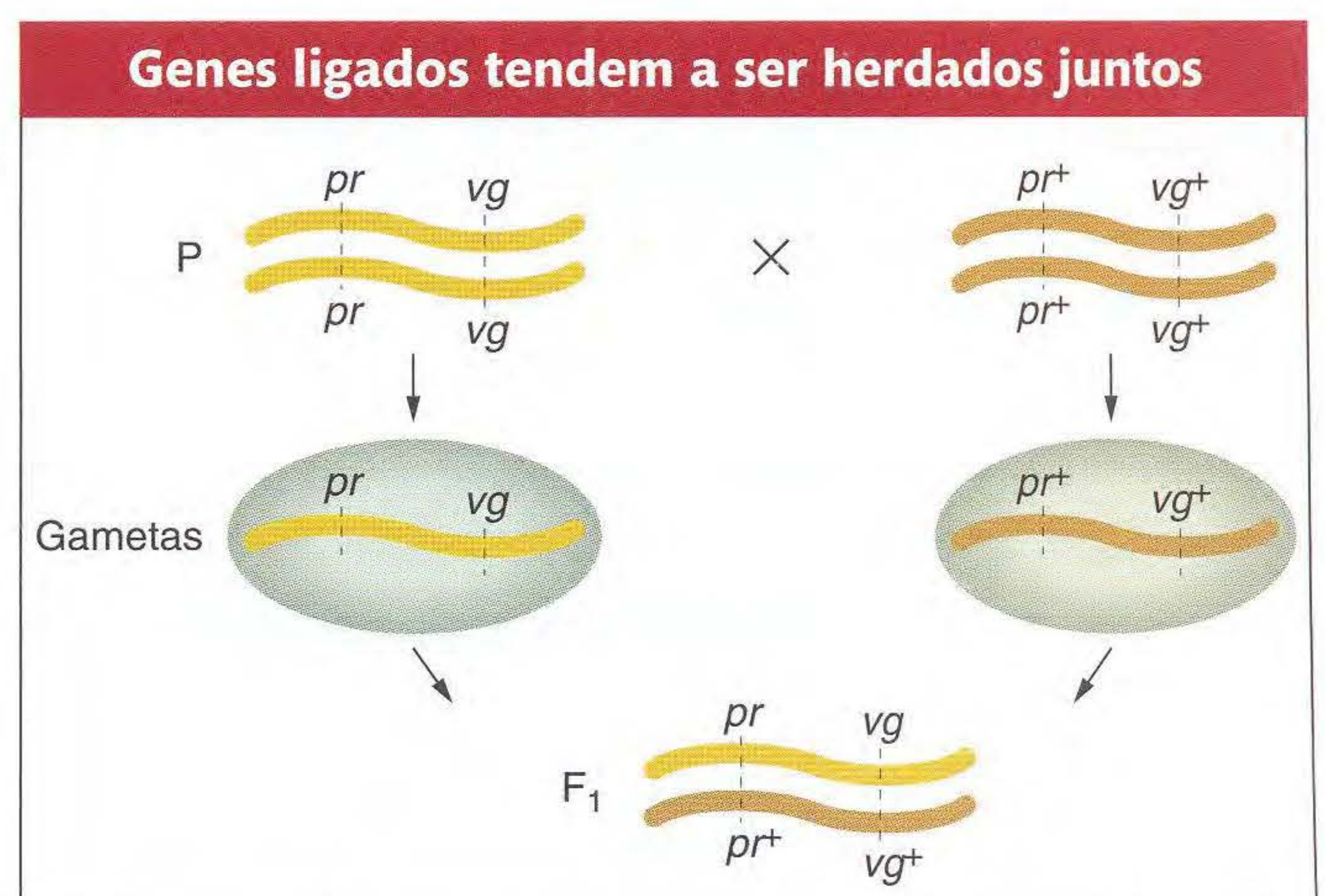


FIG. 4.2 Herança simples de dois genes situados no mesmo par de cromossomos. Os mesmos genes estão presentes juntos em um cromossomo em ambos os genitores e na prole.

Vejamos agora um outro cruzamento que Morgan fez com o uso dos mesmos alelos, mas em combinação diferente. Nesse cruzamento, cada genitor é homozigoto para o alelo tipo selvagem de um gene e o alelo mutante do outro. Novamente, as fêmeas F_1 foram submetidas a cruzamento-teste:

$$P \quad pr^+/pr^+ \cdot vg/vg \times pr/pr \cdot vg^+/vg^+$$

$$\text{Gametas} \quad pr^+ \cdot vg \quad pr \cdot vg^+$$

$$\text{Diíbrido de } F_1 \quad pr^+/pr \cdot vg^+/vg$$

Cruzamento-teste:

$$\begin{array}{ccc} \text{♀ } pr^+/pr \cdot vg^+/vg & \times & \text{♂ } pr/pr \cdot vg/vg \\ \text{Fêmea diíbrida de } F_1 & & \text{Macho testador} \end{array}$$

A seguinte prole foi obtida do cruzamento-teste:

$pr^+ \cdot vg^+$	157
$pr \cdot vg$	146
$pr^+ \cdot vg$	965
$pr \cdot vg^+$	1.067
	2.335

Novamente, esses resultados não são nem próximos da proporção mendeliana 1:1:1:1. Agora, entretanto, as classes recombinantes são o contrário das da primeira análise, $pr^+ vg^+$ e $pr vg$. Mas note que sua frequência é aproximadamente a mesma: $(157 + 146)/2.335 \times 100 = 12,9\%$. Novamente, é sugerida a ligação, mas, nesse caso, o diíbrido de F_1 deve ter sido o seguinte:

$$\frac{pr^+ \quad vg}{pr \quad vg^+}$$

Os resultados do cruzamento-teste diíbrido são como os comumente encontrados em genética. Eles seguem o padrão geral:

Duas classes não-recombinantes igualmente freqüentes somando um mais do que 50%

Duas classes recombinantes igualmente freqüentes totalizando menos de 50%

Quando dois genes estão muito próximos no mesmo par de cromossomos (isto é, estão ligados), eles não se distribuem independentemente, mas produzem uma frequência de recombinantes menor que 50%. Assim, contrariamente, a frequência de recombinação menor que 50% é diagnóstica de ligação.

Como os crossings produzem recombinantes para genes ligados

A hipótese de ligação explica por que as recombinações alélicas das gerações parentais permanecem juntas: os genes são fisicamente ligados pelo segmento de cromossomo entre eles. Mas exatamente de que modo *quaisquer* recombinantes são produzidos quando os genes estão ligados? Morgan sugeriu

O crossing over produz novas combinações alélicas

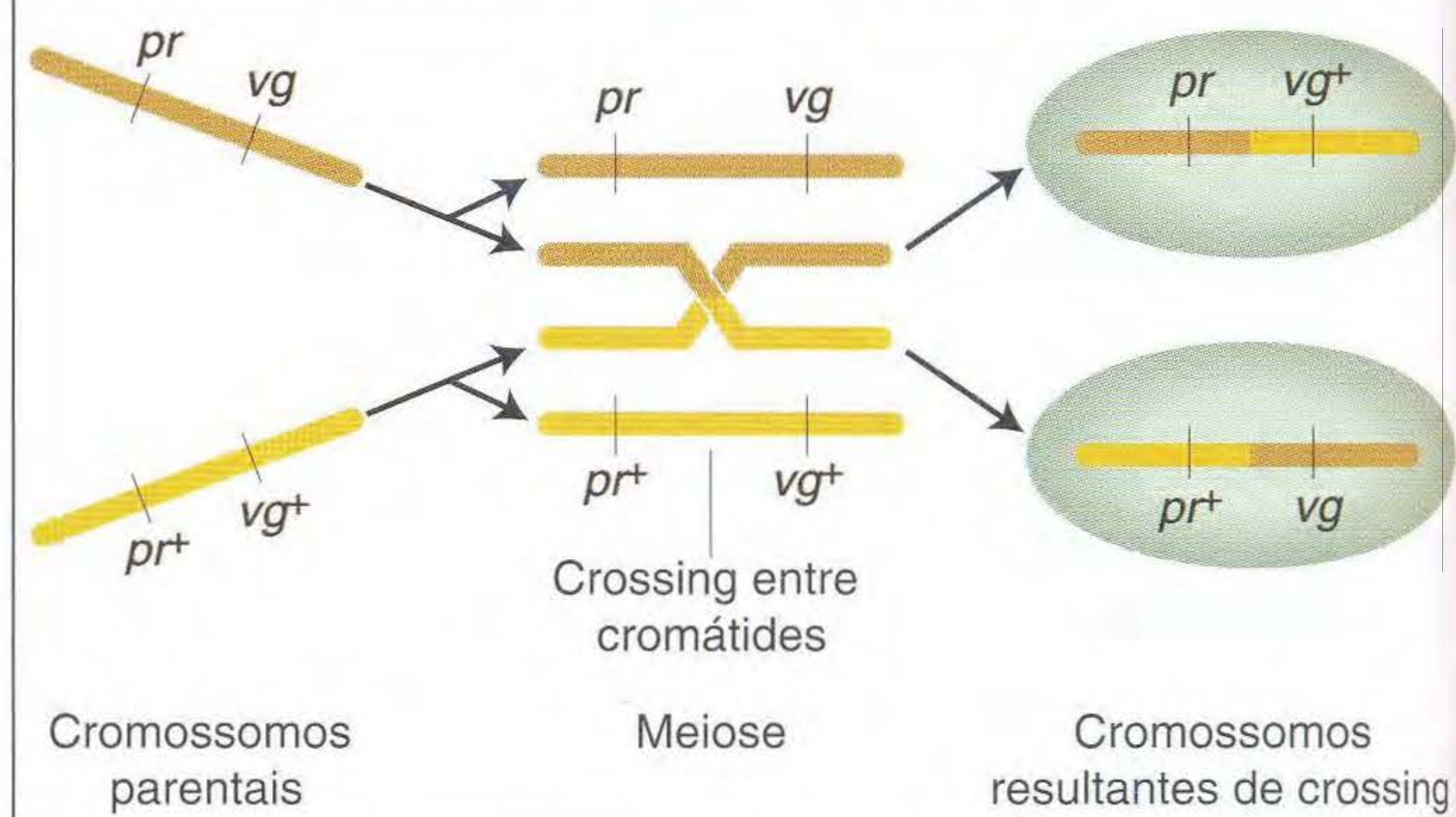


FIG. 4.3 A troca de partes pelo crossing pode produzir cromossomos gaméticos cujas combinações alélicas diferem das combinações parentais.

que, quando cromossomos homólogos fazem par na meiose, os cromossomos ocasionalmente se rompem e trocam partes em um processo chamado de **crossing over**. A Figura 4.3 ilustra essa troca física de segmentos cromossômicos. As duas novas combinações são chamadas de **produtos de crossing**.

Existe algum processo microscopicamente observável que possa explicar o crossing over? Na meiose, quando cromossomos duplicados formam pares — em termos genéticos, quando duas díades se unem como um bivalente — geralmente se forma uma estrutura em forma de cruz, chamada *quiasma*, entre as duas cromátides não-irmãs. Os quiasmas são mostrados na Figura 4.4. Para Morgan, o aparecimento dos quiasmas confirmava visualmente o conceito de crossing over. (Note que os quiasmas parecem indicar que as *cromátides*, não os cromossomos duplicados, participam de um crossing. Voltaremos a esse ponto mais adiante.)

Quiasmas são os locais de crossing

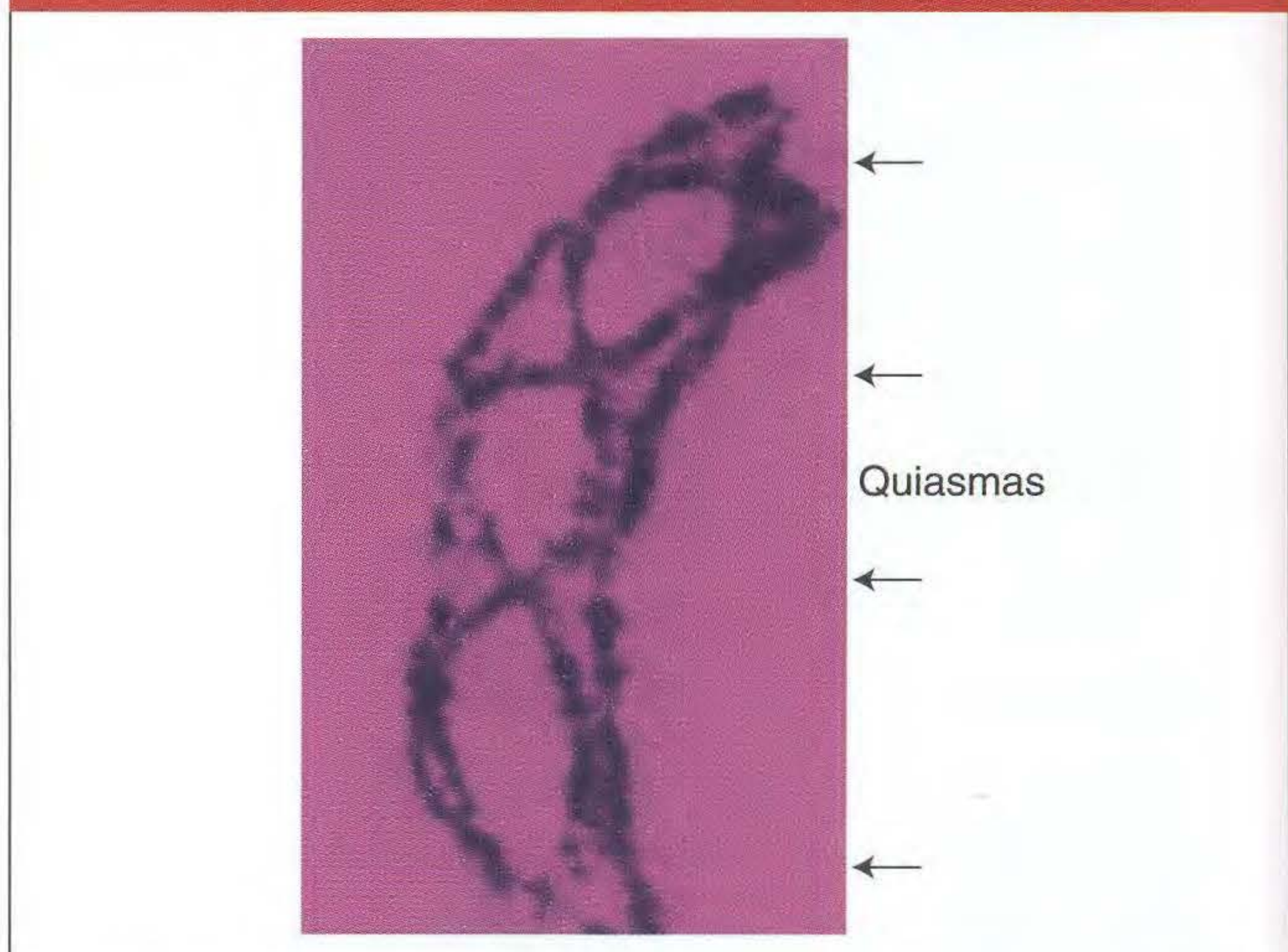


FIG. 4.4 Vários quiasmas aparecem nessa foto tirada durante a meiose em testículo de gafanhoto. [John Cabisco/Visuals Unlimited.]

Para genes ligados, os recombinantes são produzidos por crossings. Os quiasmas são as manifestações visíveis dos crossings.

Simbolismo de ligação e terminologia

O trabalho de Morgan mostrou que os genes ligados em um diíbrido podem estar presentes em uma das duas conformações básicas. Em uma, os dois alelos dominantes, ou tipo selvagem, estão presentes no mesmo homólogo (como na Figura 4.3; esse arranjo é chamado **conformação cis** (cis significa adjacente). Na outra, eles estão em homólogos diferentes, no que é chamado de **conformação trans** (trans significa oposta). As duas conformações são escritas do seguinte modo:

Cis	AB/ab ou $+/+$
Trans	Ab/aB ou $+b/a+$

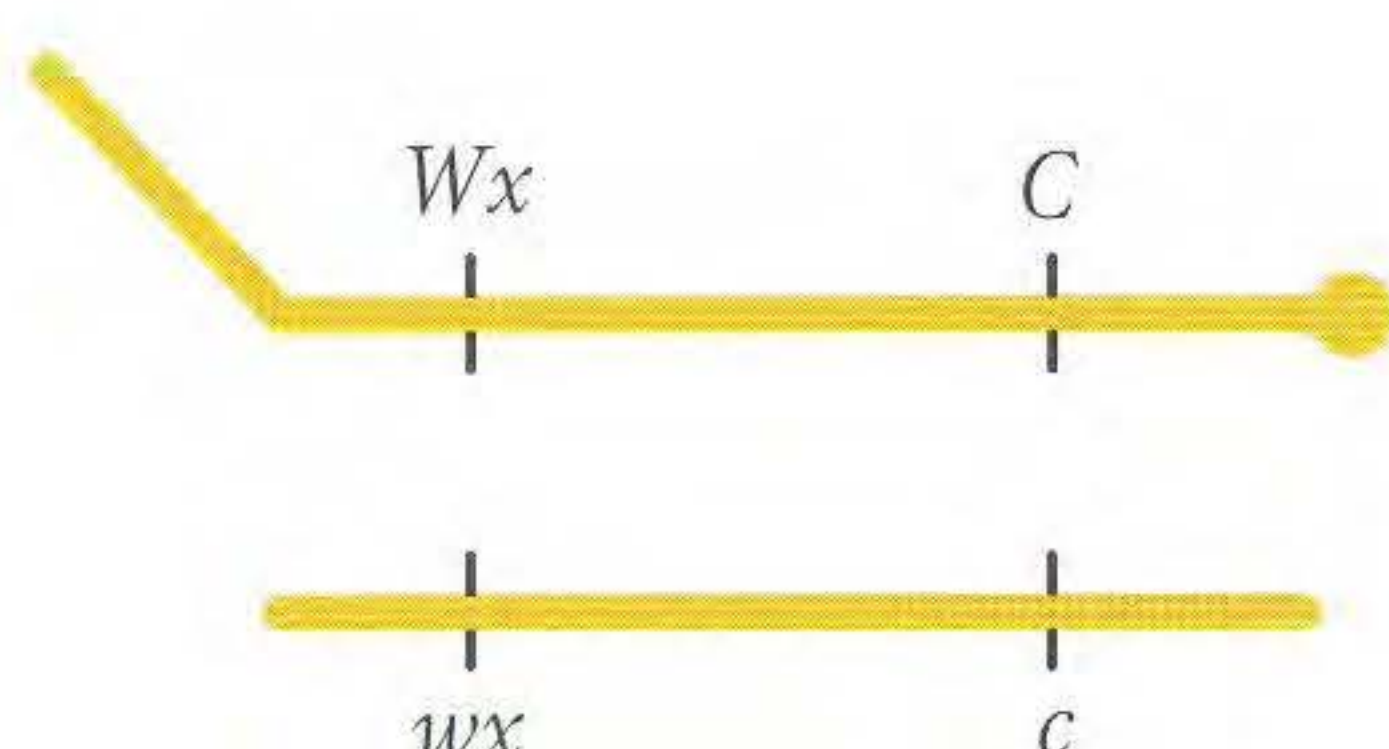
Note as seguintes convenções que pertencem ao simbolismo de ligação.

1. Alelos no mesmo homólogo não têm pontuação entre si.
2. Uma barra simbolicamente separa os dois homólogos.
3. Os alelos são sempre escritos na mesma ordem em cada homólogo.
4. Como em capítulos anteriores, os genes conhecidos como estando em cromossomos diferentes (genes não-ligados) são mostrados separados por um ponto-e-vírgula, por exemplo, $A/a ; C/c$.
5. Neste livro, os genes de ligação *desconhecida* são mostrados separados por um ponto, $A/a \cdot D/d$.

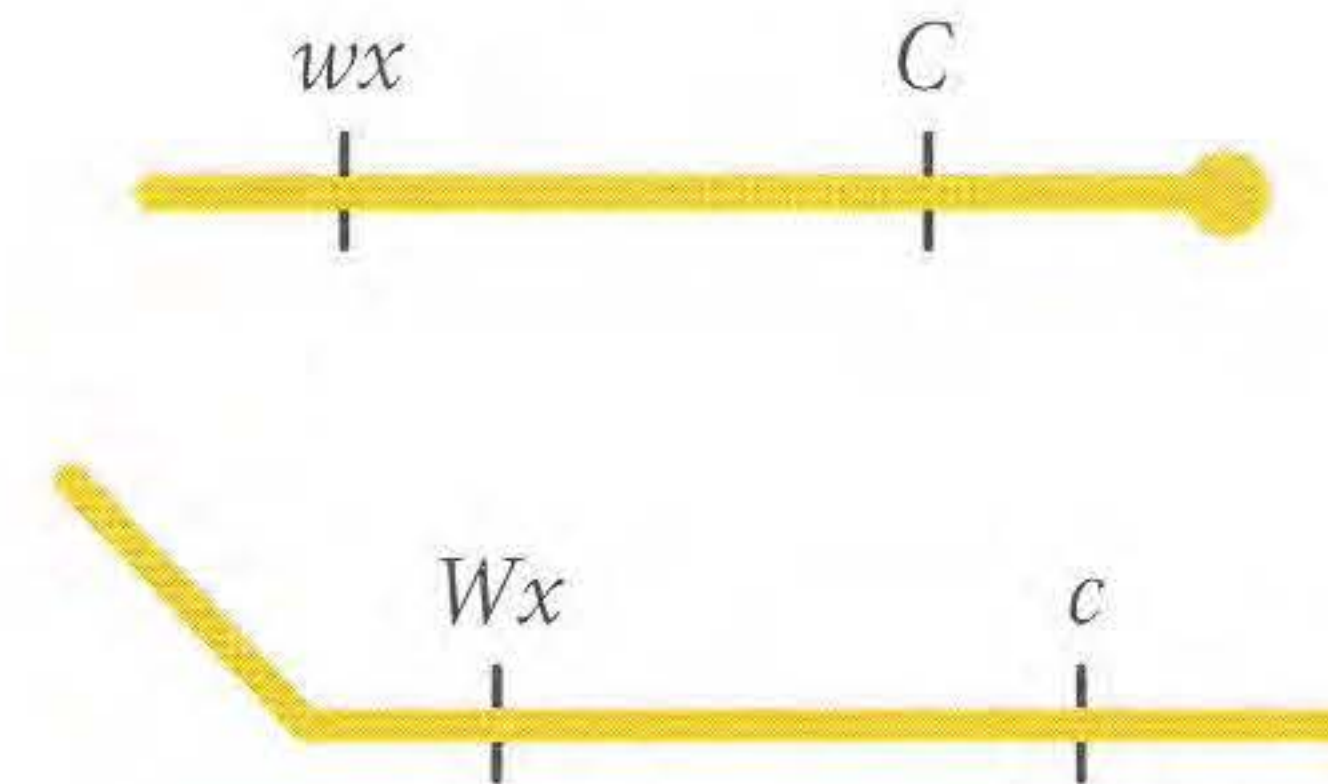
Evidência de que o crossing over é um processo de quebra e reunião

A idéia de que os recombinantes são produzidos por algum tipo de troca de material entre cromossomos homólogos era tentadora. Mas era necessário experimentação para testar essa hipótese. Uma primeira etapa era encontrar um caso no qual a troca de partes entre os cromossomos seria visível ao microscópio. Vários pesquisadores abordaram esse problema do mesmo modo, e uma de suas análises é a seguinte.

Em 1931, Harriet Creighton e Barbara McClintock estavam estudando dois genes de milho que eles sabiam estarem situados no cromossomo 9. Um afetava a cor da semente (C , colorido; c , incolor), e o outro afetava a composição do endosperma (Wx , *waxy*; wx , amiláceo). A planta era um diíbrido em conformação cis. Entretanto, em uma planta, o cromossomo 9 levando os alelos C e Wx era incomum pelo fato de possuir um elemento grande, densamente colorido (chamado *knob*) na ponta C e um trecho maior do cromossomo na ponta Wx ; logo, o heterozigoto era



Na prole de um cruzamento-teste dessa planta, eles compararam recombinantes e genótipos parentais. Eles descobriram que todos os recombinantes herdaram um ou o outro dos seguintes cromossomos, dependendo de sua constituição recombinante:



Assim, havia uma correlação precisa entre o evento *genético* e o surgimento de recombinantes e o evento *cromossômico* do crossing over. Conseqüentemente, os quiasmas pareciam ser os sítios de troca, embora o que foi considerado como sendo um teste definitivo só foi feito em 1978.

O que podemos dizer sobre o mecanismo molecular de troca cromossômica em um crossing? A resposta simples é que o crossing resulta da quebra e reunião do DNA. Duas quebras cromossômicas parentais ocorrem na mesma posição, e então cada pedaço junta-se ao pedaço vizinho de *outro* cromossomo. No Capítulo 14, estudaremos os modelos dos processos moleculares que permitem ao DNA quebrar-se e reunir-se de um modo preciso a não haver ganho nem perda de material genético.

Um crossing é a quebra de duas moléculas de DNA na mesma posição e sua reunião em duas combinações recombinantes recíprocas.

Evidência de que o crossing over ocorre no estágio de quatro cromátides

Como já foi notado, a representação diagramática do crossing over na Figura 4.3 mostra o crossing ocorrendo no estágio de quatro cromátides na meiose; em outras palavras, os crossings ocorrem entre *cromátides* não-irmãs. Entretanto, era *teoricamente* possível que o crossing over ocorresse antes da replicação, no estágio de *dois* cromossomos. Essa incerteza foi resolvida pela análise genética dos organismos cujos quatro produtos da meiose permaneciam juntos em grupos de quatro chamados *tétrades*. Esses organismos, que vimos nos Capítulos 2 e 3, são fungos e algas unicelulares. Os produtos da meiose de uma única tétrade podem ser isolados, o que é equivalente a isolar todas as quatro cromátides de um único meiócito. A análise de tétrades dos cruzamentos *nos quais os genes estão ligados* mostra muitas tétrades que contêm quatro combinações alélicas diferentes. Por exemplo, do cruzamento

$$AB \times ab$$

algumas (mas não todas) as tétrades contêm quatro genótipos:

AB
 Ab
 aB
 ab

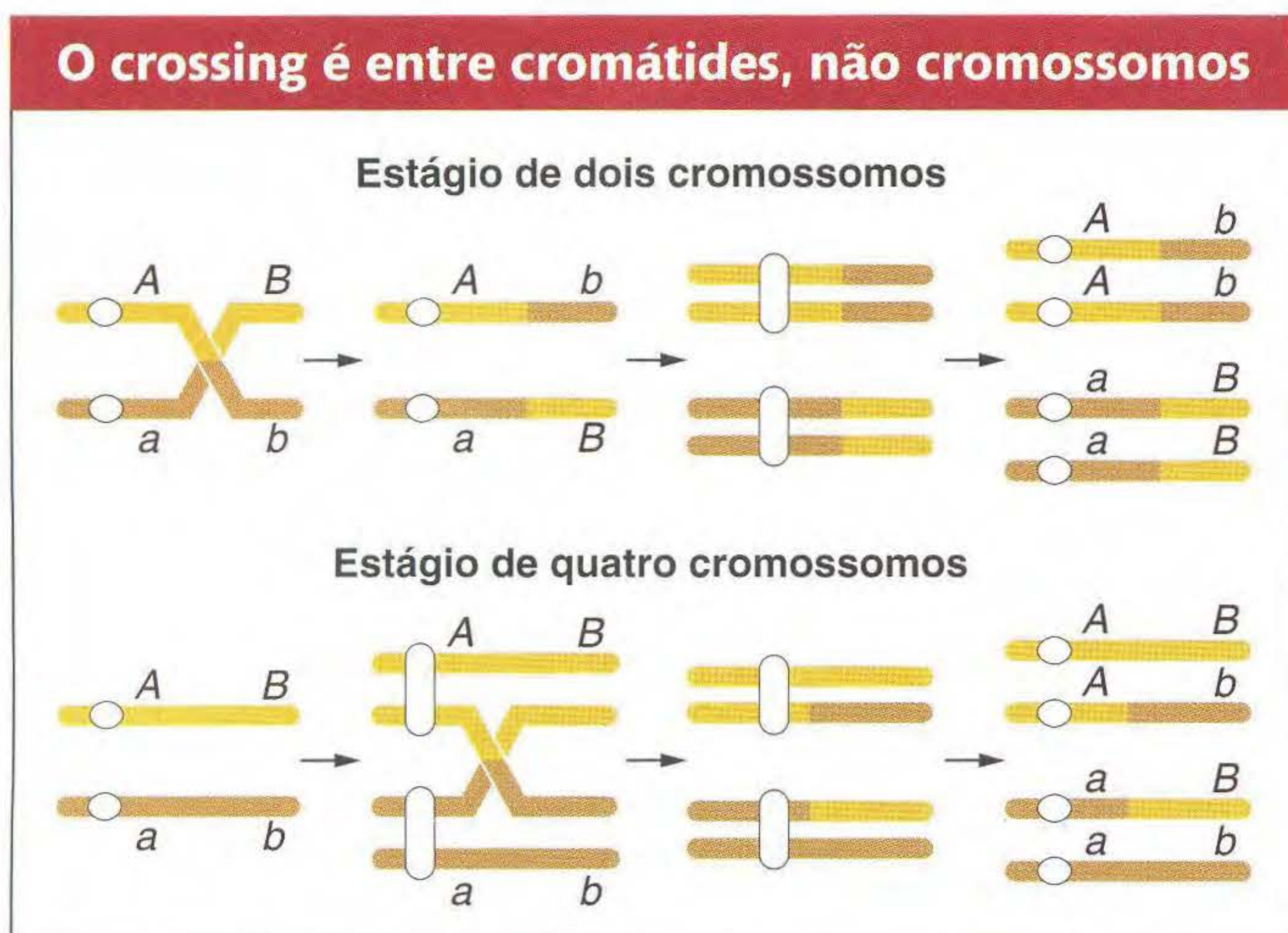


FIG. 4.5 O crossing over ocorre no estágio de quatro cromátides. Como mais de dois produtos diferentes de uma única meiose podem ser vistos em algumas tétrades, o crossing over não pode ocorrer no estágio de dois filamentos (antes da replicação do DNA). O círculo branco indica a posição do centrômero. Quando as cromátides-irmãs são visíveis, o centrômero aparece não-replicado.

Esse resultado só pode ser explicado se os crossings ocorrerem no estágio de quatro cromátides, pois, se os crossings ocorrerem no estágio de duas cromátides, só pode haver um máximo de dois genótipos diferentes em uma tétrade individual. Esse raciocínio é ilustrado na Figura 4.5.

Múltiplos crossings podem incluir mais de duas cromátides

A análise de tétrades também pode mostrar duas outras características importantes do crossing over. Primeira, em alguns meiócitos individuais, vários crossings podem ocorrer ao longo de um par de cromossomos. Segundo, em qualquer meiócito, esses vários crossings podem trocar material entre mais de duas cromátides. Para pensar sobre isso, precisamos observar o caso mais simples: crossings duplos. Para estudar crossings duplos, precisamos de três genes ligados. Por exemplo, se três loci estão ligados em um cruzamento tal como

$$ABC \times abc$$

são possíveis muitos tipos diferentes de tétrades, mas alguns tipos são informativos na presente conexão porque podem responder por apenas crossings duplos nos quais mais de duas cromátides tomam parte. Considere a seguinte tétrade como um exemplo:

ABc
AbC
aBC
abc

Essa tétrade deve ser explicada por dois crossings nos quais *três* cromátides tomam parte, como mostrado na Figura 4.6a. Além disso, o seguinte tipo de tétrade mostra que todas as quatro cromátides podem participar do crossing over na mesma meiose (Figura 4.6b):

ABc
Abc
aBC
abC

Portanto, para qualquer par de cromossomos homólogos, duas, três ou quatro cromátides podem tomar parte em eventos de crossing em um único meiócito. Note, entretanto, que qualquer crossing está entre duas cromátides.

Você deve estar pensando em crossings entre cromátides-irmãs. Eles ocorrem, mas são raros. Eles não produzem novas combinações alélicas e, portanto, não são geralmente considerados.

► 4.2 Mapeamento por Frequência de Recombinação

A frequência de recombinantes produzida pelo crossing over é a chave do mapeamento cromossômico. A análise de tétrade de fungos mostrou que, para qualquer dois genes ligados específicos, os crossings ocorrem entre eles em alguns meiócitos, mas não em todos (Figura 4:7). Quanto mais distantes os genes estiverem, mais provavelmente ocorrerá um crossing e maior a proporção de produtos recombinantes. Assim, a proporção de recombinantes é a chave para a distância que separa dois loci gênicos em um mapa cromossômico.

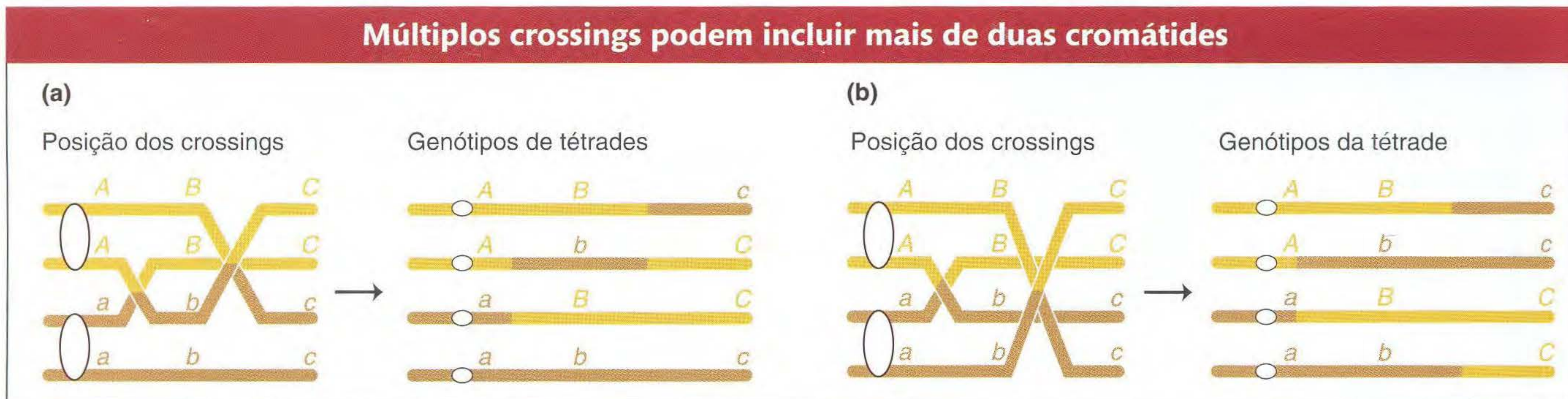


FIG. 4.6 Crossings duplos podem incluir (a) três cromátides ou (b) quatro cromátides.

Recombinação
meiótica entre
genes ligados
por crossing
over

W W W . ANIMAÇÃO COMPUTADORIZADA

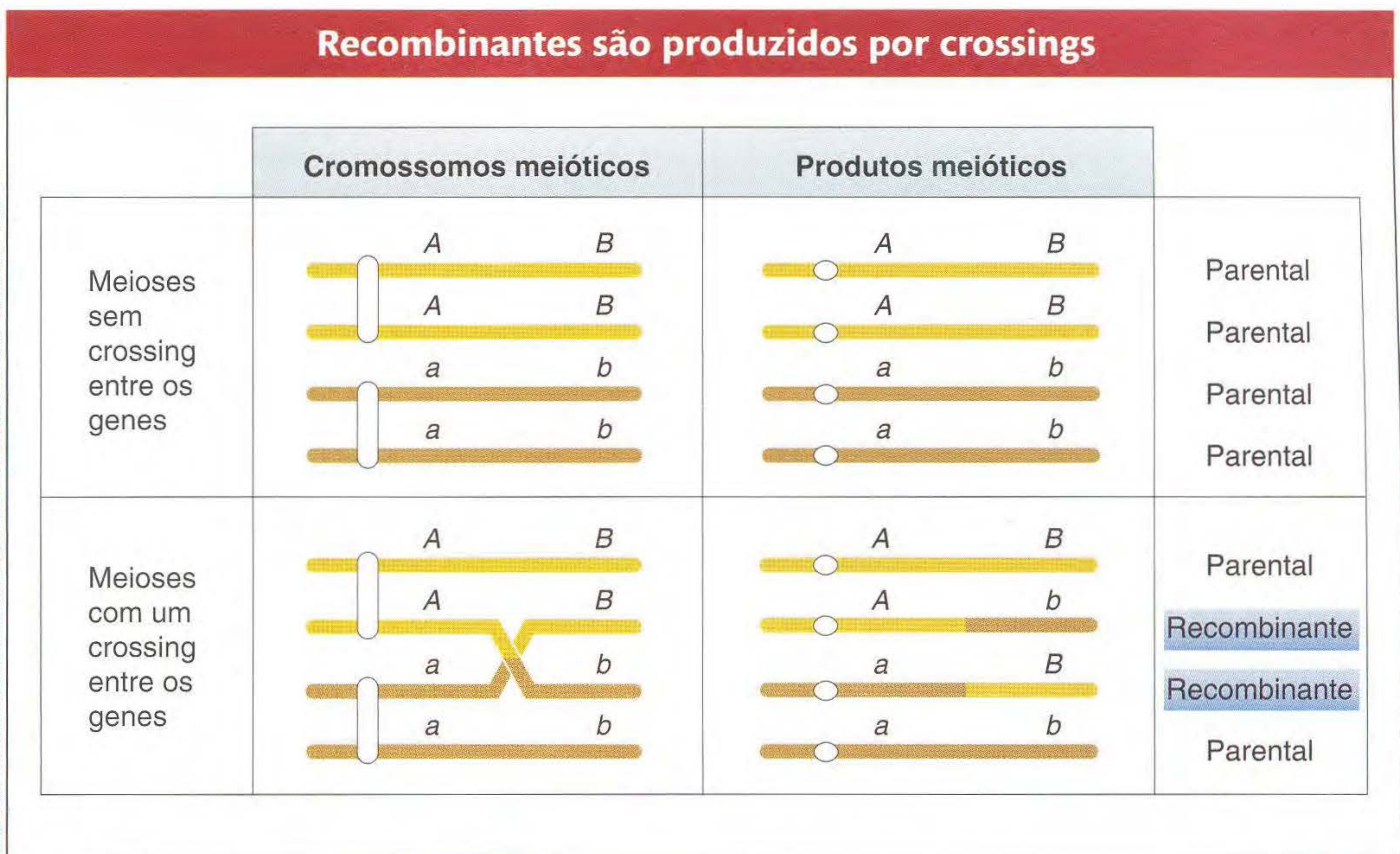


FIG. 4.7 Os recombinantes surgem das meioses nas quais ocorre o crossing entre cromátides não-irmãs.



Como já foi dito com relação aos dados de Morgan, a frequência de recombinantes foi significativamente menor do que 50%, especificamente 10,7%. A Figura 4.8 mostra a situação geral para ligação na qual os recombinantes são menos que 50%. As frequências de recombinantes para diferentes genes ligados variam de 0 a 50%, dependendo de sua proximidade (veja mais adiante). Quanto mais distantes estão os genes, mais proximamente suas frequências de recombinantes aproximam-se de 50%, e, em tais casos, não podemos decidir se os genes estão ligados ou estão em cromossomos diferentes. E quanto a frequências de recombinantes maiores que 50%? A resposta é que tais frequências *nunca* são observadas, como será provado mais adiante.

Note, na Figura 4.7, que um único crossing gera dois produtos recombinantes recíprocos, que explicam por que as classes recombinantes são, em geral, aproximadamente iguais em frequência. O corolário desse ponto é que dois tipos não-recombinantes parentais também devem ser iguais em frequência, como também observado por Morgan.

Unidades de mapa

O método básico do mapeamento de genes com o uso de frequências de recombinantes foi desenvolvido por um estudante de Morgan. À medida que Morgan estudava mais e mais os genes ligados, ele percebeu que a proporção de prole recombinante variava consideravelmente, dependendo de quais genes ligados estavam sendo estudados, e ele pensou que tal variação na frequência de recombinantes podia indicar de algum modo as distâncias reais que separam os genes nos cromossomos. Morgan atribuiu a quantificação desse processo a um estudante de graduação, Alfred Sturtevant, que também se tornou um grande geneticista. Morgan pediu a Sturtevant que tentasse dar alguns sentidos aos dados de crossing over entre genes ligados diferentes. Em uma tarde,

Sturtevant desenvolveu um método para mapear genes que ainda hoje é usado. Nas próprias palavras de Sturtevant, “Na parte final de 1911, conversando com Morgan, subitamente

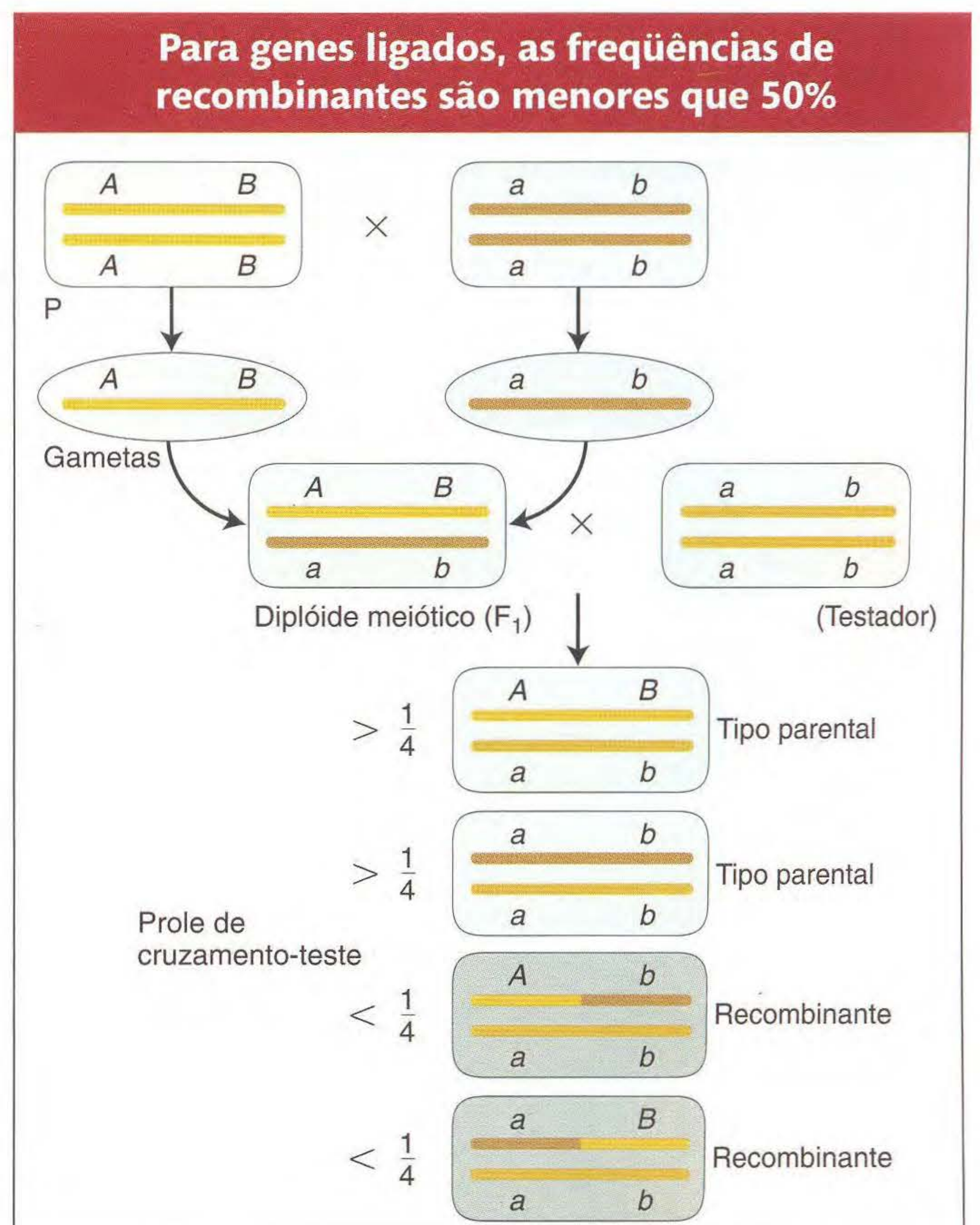


FIG. 4.8 Um cruzamento-teste revela que as frequências de recombinantes que surgem dos crossings entre genes ligados são menores que 50%.

percebi que as variações na força de ligação, já atribuída por Morgan a diferenças na separação espacial dos genes, oferecia a possibilidade de determinar as seqüências na dimensão linear de um cromossomo. Fui para casa e passei a maior parte da noite (ele não fez a tarefa de graduação) produzindo o primeiro mapa cromossômico.”

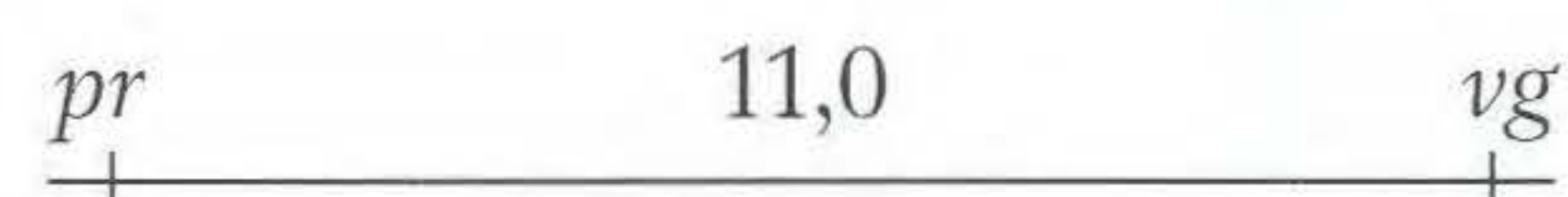
Como um exemplo da lógica de Sturtevant, considere os resultados do cruzamento-teste de Morgan com os genes *pr* e *vg*, dos quais ele calculou uma freqüência de recombinantes de 10,7%. Sturtevant sugeriu que podemos usar essa percentagem de recombinantes como um índice quantitativo da distância linear entre dois genes em um mapa genético, ou **mapa de ligação**, como às vezes é chamado.

A idéia básica aqui é bem simples. Imagine dois genes específicos posicionados a certa distância fixa. Agora, imagine um crossing over aleatório ao longo dos homólogos pareados. Em algumas meioses, as cromátides não-irmãs fazem crossing ao acaso na região cromossômica entre esses genes. Dessas meioses, são produzidos recombinantes. Em outras divisões meióticas, não há crossings entre esses genes. Nenhum recombinante resulta dessas meioses. (Volte à Figura 4.7 para uma ilustração diagramática.) Sturtevant postulou uma proporcionalidade: quanto maior a distância entre genes ligados, maior a chance de crossings na região entre os genes e, portanto, maior a proporção de recombinantes que seriam produzidas. Assim, determinando a freqüência de recombinantes, podemos obter uma medida da distância de mapa entre os genes. De fato, Sturtevant definiu uma **unidade de mapa genético (u.m.)** como a distância entre genes para a qual um produto de meiose em 100 é recombinante. Por exemplo, a **freqüência de recombinantes (FR)** de 10,7 obtida por Morgan é definida como 10,7 u.m. Uma unidade

de mapa às vezes é chamada de um **centimorgan (cM)** em homenagem a Thomas Hunt Morgan.

Esse método produz um mapa linear correspondente à linearidade cromossômica? Sturtevant previu que, em um mapa linear, se 5 unidades de mapa (5 u.m.) separam os genes *A* e *B*, enquanto 3 u.m. separam os genes *A* e *C*, então a distância que separa *B* e *C* deve ser ou 8 ou 2 u.m. (Figura 4.9). Sturtevant viu que sua previsão era correta. Em outras palavras, sua análise sugeriu fortemente que os genes estão dispostos em alguma ordem linear, tornando as distâncias de mapa aditivas. (Existem algumas pequenas mas não insignificantes exceções, como veremos mais adiante.)

Como é representado um mapa? Como um exemplo, em *Drosophila*, o locus do gene de cor de olho e o locus do gene do tamanho da asa distam aproximadamente 11 u.m., como já mencionado. A relação geralmente é diagramada do seguinte modo:



Em geral, referimo-nos ao locus do gene de cor de olho abreviadamente como o “locus *pr*”, devido ao primeiro alelo mutante descoberto, mas queremos dizer o lugar no cromossomo onde *qualquer* alelo desse gene será encontrado, mutante ou tipo selvagem.

Como dito nos Capítulos 2 e 3, a análise genética pode ser aplicada em duas direções opostas. Esse princípio é aplicável a freqüências de recombinantes. Em uma direção, as freqüências de recombinantes podem ser usadas para fazer mapas. Na outra direção, quando determinado mapa é estabelecido com unidades de mapa, podemos prever as freqüências da prole em classes diferentes. Por exemplo, a distância genética

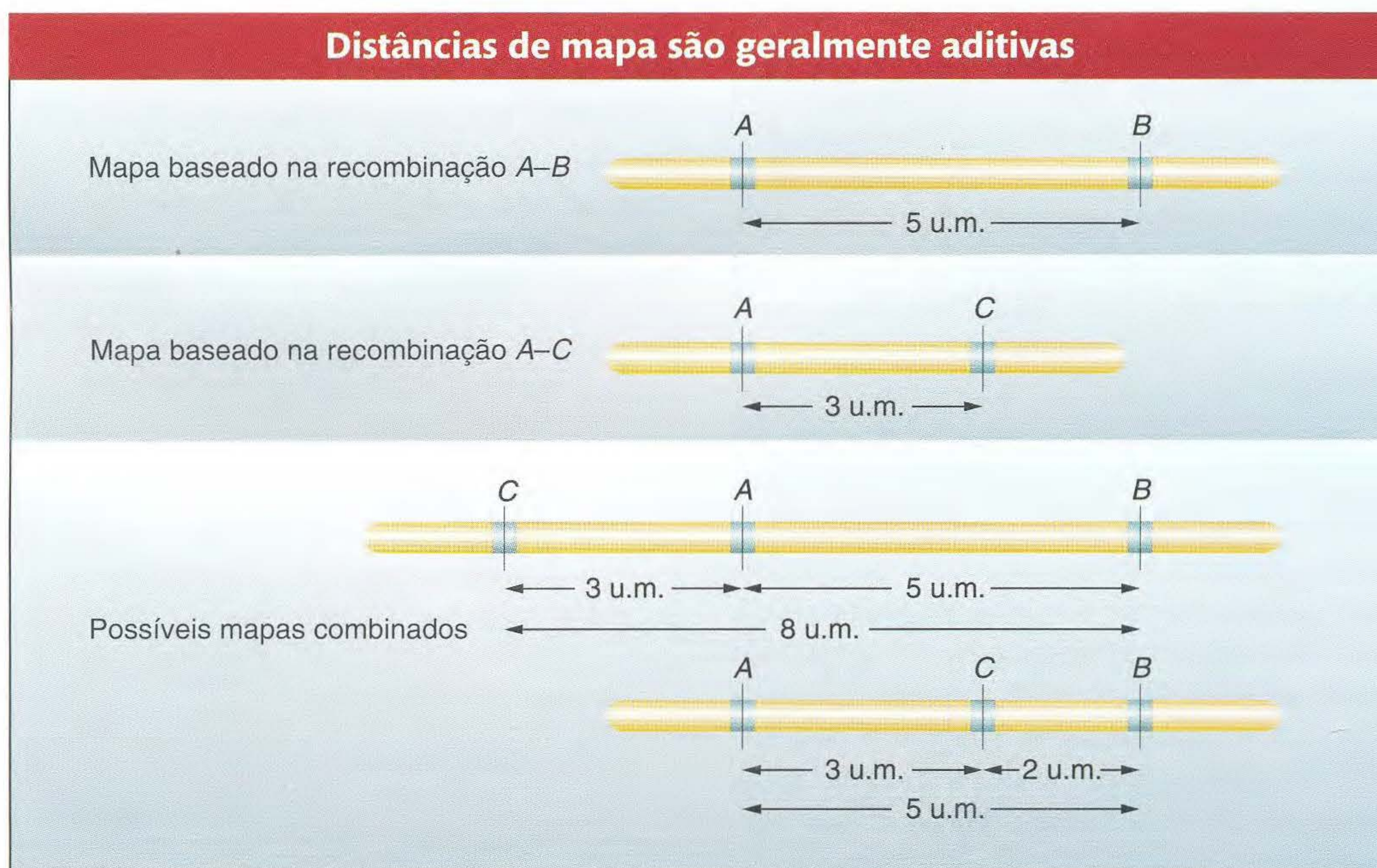


FIG. 4.9 Uma região cromossômica contendo três genes ligados. Como as distâncias de mapa são aditivas, o cálculo das distâncias A-B e A-C nos deixa com duas possibilidades mostradas para a distância B-C.

entre os loci *pr* e *vg* em *Drosophila* é de aproximadamente 11 unidades de mapa. Assim, conhecendo esse valor, sabemos que haverá 11% de recombinantes na prole de um cruzamento-teste de uma fêmea diíbrida heterozigota em conformação cis (*pr vg/pr⁺ vg⁺*). Esses recombinantes consistirão em dois recombinantes recíprocos de igual frequência: assim, 5,5% serão *pr vg⁺* e 5,5% serão *pr⁺ vg*. Também sabemos que 100 - 11 = 89% serão não-recombinantes em duas classes iguais, 44,5% *pr⁺ vg⁺* e 44,5% *pr vg*. (Note que a contribuição do testador *pr vg* foi ignorada ao se escrever esses genótipos.)

Há uma forte implicação em que a “distância” em um mapa de ligação é uma distância física ao longo de um cromossomo, e Morgan e Sturtevant certamente queriam dizer exatamente isso. Mas devemos perceber que o mapa de ligação é uma entidade *hipotética* construída de uma análise puramente genética. O mapa de ligação pode ter sido obtido sem mesmo se conhecer que os cromossomos existiam. Além disso, neste ponto em nossa discussão, não podemos dizer se as “distâncias genéticas” calculadas por meio de frequências recombinantes representam de algum modo as distâncias físicas reais nos cromossomos. Entretanto, os mapas físicos mostraram que as distâncias genéticas são, de fato, mais ou menos proporcionais a distâncias baseadas em recombinação. Existem exceções causadas por pontos quentes de recombinação, locais no genoma onde ocorre crossing over mais frequentemente que o usual. A presença de pontos quentes causa expansão proporcional de algumas regiões do mapa. Os bloqueios de recombinação, que têm o efeito oposto, também são conhecidos.

Um resumo do modo pelo qual os recombinantes do crossing over são usados no mapeamento é mostrado na Figura 4.10. Os crossings ocorrem mais ou menos aleatoriamente ao longo do par cromossômico. Em geral, em regiões mais longas, o número médio de crossings é maior, e, concordantemente, os recombinantes são mais frequentemente obtidos, traduzindo-se em uma distância de mapa mais longa.

A recombinação entre genes ligados pode ser usada para mapear sua distância em um cromossomo. A unidade de mapeamento (1 u.m.) é definida como uma frequência de recombinação de 1%.

Cruzamento-teste de três pontos

Até agora, vimos linhagens em cruzamentos de diíbridos (duplos heterozigotos) com testadores duplos recessivos. O nível seguinte de complexidade é um cruzamento de um triíbrido (tripo heterozigoto) com um testador triplo recessivo. Esse tipo de cruzamento, chamado de **cruzamento-teste de três pontos**, é comumente usado na análise de ligação. A meta é deduzir se os três genes estão ligados e, se estiverem, deduzir sua ordem e as distâncias de mapa entre eles.

Veamos um exemplo, também de *Drosophila*. Em nosso exemplo, os alelos mutantes são *v* (olhos vermillion), *cv* (*crossveinless*, ou ausência de nervuras nas asas, e *ct* (*cut*, bordas aparadas das asas). A análise é feita fazendo os seguintes cruzamentos:

$$\begin{array}{lcl} \text{P} & v^+/v^+ \cdot cv/cv \cdot ct/ct & \times \quad v/v \cdot cv^+/cv^+ \cdot ct^+/ct^+ \\ & \downarrow & \\ \text{Gametas} & v^+ \cdot cv \cdot ct & v \cdot cv^+ \cdot ct^+ \\ \text{Triíbrido de } F_1 & v/v^+ \cdot cv/cv^+ \cdot ct/ct^+ & \end{array}$$

As fêmeas triíbridas são submetidas a cruzamentos-teste com machos triplos recessivos:

$$\begin{array}{lcl} \text{♀ } v/v^+ \cdot cv/cv^+ \cdot ct/ct^+ & \times & \text{♂ } v/v \cdot cv/cv \cdot ct/ct \\ \text{Fêmea triíbrida de } F_1 & & \text{Macho testador} \end{array}$$

De qualquer triíbrido, apenas $2 \times 2 \times 2 = 8$ genótipos de gameta são possíveis. Eles são os genótipos vistos na prole do cruzamento-teste. O quadro seguinte mostra o número de cada um dos oito genótipos gaméticos em uma amostra de 1.448 moscas da prole. As colunas mostram quais genótipos são recombinantes (R) para os loci tomados dois de cada vez. Devemos ser cuidadosos em nossa classificação dos tipos parentais e recombinantes. Note que os genótipos parentais que entram para os triplos heterozigotos são $v^+ \cdot cv \cdot ct$ e $v \cdot cv^+ \cdot ct^+$. Qualquer outra combinação além dessas duas constituem um recombinante.

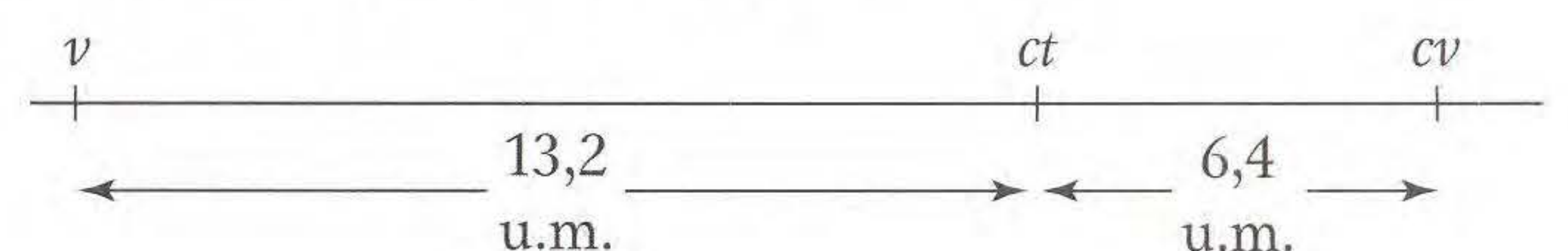
Gametas	Recombinante para os loci		
	<i>v</i> e <i>cv</i>	<i>v</i> e <i>ct</i>	<i>cv</i> e <i>ct</i>
$v \cdot cv^+ \cdot ct^+$	580		
$v^+ \cdot cv \cdot ct$	592		
$v \cdot cv \cdot ct^+$	45	R	
$v^+ \cdot cv^+ \cdot ct$	40	R	
$v \cdot cv \cdot ct$	89	R	R
$v^+ \cdot cv^+ \cdot ct^+$	94	R	R
$v \cdot cv^+ \cdot ct$	3		R
$v^+ \cdot cv \cdot ct^+$	5		R
	1.448	268	191
			93

Vamos analisar os loci dois de cada vez, começando com os loci *v* e *cv*. Em outras palavras, vamos olhar apenas as duas primeiras colunas sob “Recombinantes para os loci” e cobrir a terceira. Como os parentais para esse par de loci são $v^+ \cdot cv$ e $v \cdot cv^+$, sabemos que os recombinantes são, por definição, $v \cdot cv$ e $v^+ \cdot cv^+$. Existem $45 + 40 + 89 + 94 = 268$ desses recombinantes. De um total de 1.448 moscas, esse número dá uma FR de 18,5%.

Para os loci *v* e *ct*, os recombinantes são $v \cdot ct$ e $v^+ \cdot ct^+$. Existem $89 + 94 + 3 + 5 = 191$ desses recombinantes entre 1.448 moscas, e, portanto, a FR = 13,2%.

Para *ct* e *cv*, os recombinantes são $cv \cdot ct^+$ e $cv^+ \cdot ct$. Existem $45 + 40 + 3 + 5 = 93$ desses recombinantes entre 1.448, e, portanto, a FR = 6,4%.

Claramente, todos os loci estão ligados, porque os valores de FR são todos consideravelmente menores que 50%. Como os loci *v* e *cv* têm o maior valor de FR, eles devem estar mais separados; portanto, o locus *ct* deve estar entre eles. Pode ser feito um mapa do seguinte modo:



Regiões maiores têm mais crossings e, assim, maiores frequências de recombinação

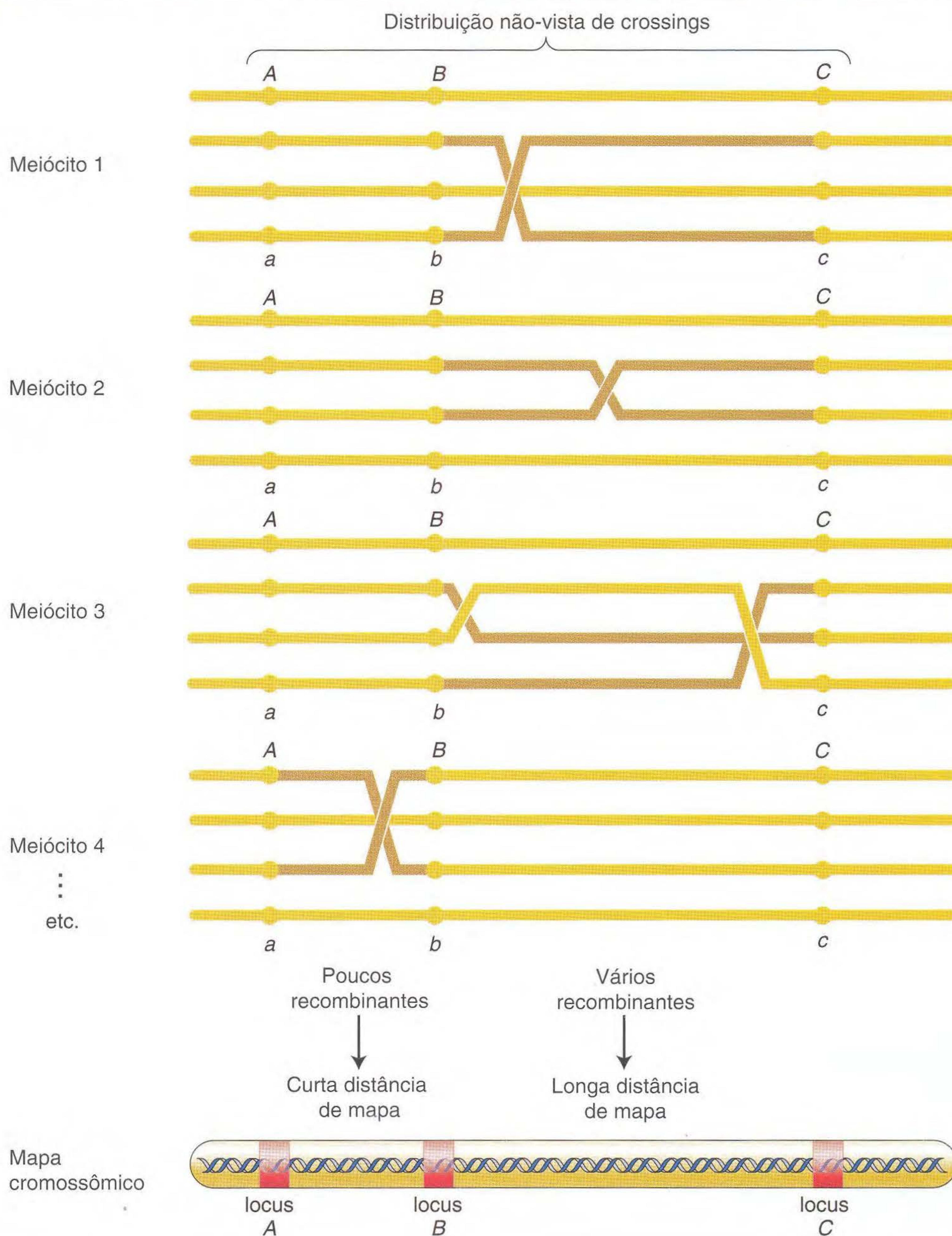


FIG. 4.10 Os crossings produzem cromátides recombinantes cujas frequências podem ser usadas para mapear genes em um cromossomo. Regiões maiores produzem mais crossings. As partes marrons mostram recombinantes para esse intervalo.

O cruzamento-teste pode ser reescrito do seguinte modo, agora que sabemos o arranjo de ligação:

$$v^+ ct cv/v ct^+ cv^+ \times v ct cv/v ct cv$$

Note vários pontos importantes aqui. Primeiro, deduzimos a ordem dos genes, que é diferente da usada em nossa lista de genótipos da prole. Como o objetivo do exercício era determinar a relação de ligação desses genes, a lista original necessariamente era arbitrária; a ordem era desconhecida antes que os dados fossem analisados. A partir disso, os genes devem ser escritos na ordem correta.

Segundo, estabelecemos definitivamente que *ct* está entre *v* e *cv*. No diagrama, colocamos arbitrariamente *v* à esquerda e *cv* à direita, mas o mapa poderia igualmente ser desenhado com esses loci invertidos.

Terceiro, note que os mapas de ligação apenas mapeiam os loci em relação uns aos outros, com o uso de unidades padrão de mapa. Não sabemos onde os loci estão em um cromossomo, nem mesmo em que cromossomo específico eles estão. Nas análises subsequentes, à medida que mais loci são mapeados em relação a esses três, o mapa cromossômico completo ficaria “concluído”.

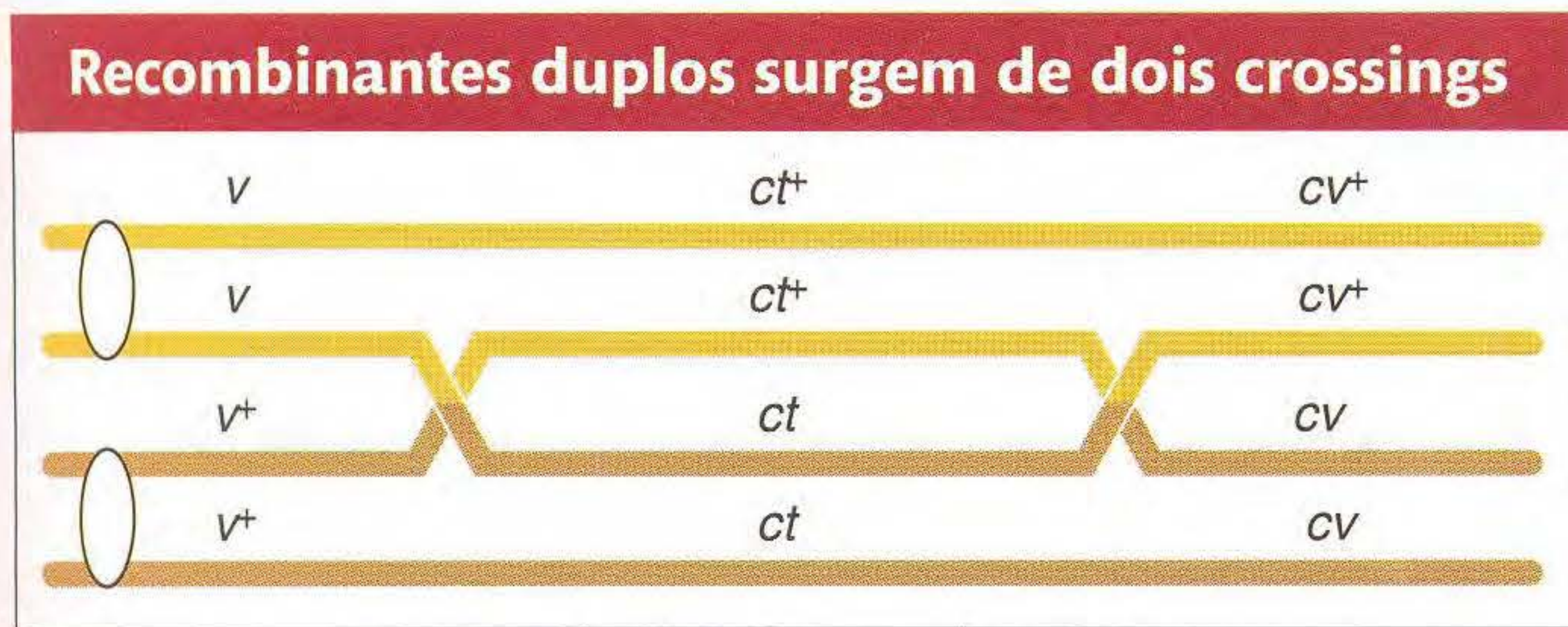


FIG. 4.11 Exemplo de um crossing duplo entre duas cromátides. Note que um crossing duplo produz cromátides de recombinantes duplos que têm combinações de alelos parentais em outros loci. A posição do centrômero não pode ser determinada pelos dados. Eles foram adicionados para complementação.

Os cruzamentos de três pontos (e mais) permitem aos geneticistas avaliar a ligação entre três (ou mais) genes e determinar a ordem dos genes, tudo em um cruzamento.

Um aspecto final a notar é que as duas menores distâncias de mapa, 13,2 u.m. e 6,4 u.m., somam 19,6 u.m., que é maior que 18,5 u.m., a distância calculada para v e cv . Por quê? A resposta a essa pergunta está no modo pelo qual tratamos as duas classes mais raras de prole (somando 8) com relação à recombinação de v e cv . Agora que temos o mapa, podemos ver que essas duas classes raras são de fato recombinantes duplos, surgindo de dois crossings (Figura 4.11). Entretanto, quando calculamos o valor FR para v e cv , não contamos os genótipos $v\ ct\ cv^+$ e $v^+\ ct^+\ cv$. Afinal, com relação a v e cv , eles são combinações parentais ($v\ cv^+$ e $v^+\ cv$). À luz de nosso mapa, entretanto, vemos que isso nos levou a subestimar a distância entre v e o locus cv . Não só deveríamos contar as duas classes mais raras, como também contar cada uma delas *duas* vezes, pois cada uma representa recombinantes duplos. Assim, podemos corrigir o valor somando os números $45 + 40 + 89 + 94 + 3 + 3 + 5 + 5 = 284$. Do total de 1.448, esse número é exatamente 19,6%, que é idêntico à soma dos dois valores componentes. (Na prática, não precisamos fazer esse cálculo, porque a soma das duas menores distâncias nos dá a melhor estimativa da distância geral.)

Dedução da ordem dos genes por inspeção

Agora que tivemos alguma experiência com o cruzamento-teste de três pontos, podemos voltar à listagem da prole e ver que, para triíbridos de genes ligados, a *ordem dos genes* pode geralmente ser deduzida por inspeção, sem uma análise da frequência de recombinação. Tipicamente, para genes ligados, temos oito genótipos nas seguintes frequências:

- dois com alta frequência
- dois com frequência intermediária
- dois com frequência intermediária diferente
- dois raros

São possíveis apenas três ordens de genes, cada uma com um gene diferente na posição média. Geralmente é verdade que as classes de recombinantes duplos são as menores, como citado por último aqui. Apenas uma ordem é compatível com as classes menores tendo sido formada por crossings duplos, como mostrado na Figura 4.12.; isto é, apenas uma ordem dá recombinantes duplos do genótipo $v\ ct\ cv^+$ e $v^+\ ct^+\ cv$. Uma regra simples para deduzir o gene do meio é que o par de alelos mudou de posição nas classes de recombinantes duplos.

Interferência

Conhecendo a existência de crossings duplos, podemos fazer perguntas sobre sua possível interdependência. Podemos perguntar: Os crossings em regiões cromossômicas adjacentes são eventos independentes ou um crossing em uma região afeta a possibilidade de haver um crossing em uma região adjacente? A resposta é que, em geral, os crossings inibem uns aos outros de modo que há uma interação chamada **interferência**. As classes de recombinantes duplos podem ser usadas para deduzir a extensão dessa interferência.

A interferência pode ser medida do seguinte modo. Se os crossings nas duas regiões são independentes, podemos usar a regra do produto (veja Capítulo 3) para prever a frequência de recombinantes duplos: essa frequência seria igual ao produto das frequências recombinantes em regiões adjacentes. Nos dados de recombinação v - ct - cv , o valor da FR v - ct é 0,132 e o valor ct - cv é 0,064; logo, não há interferência, podendo os recombinantes duplos ser esperados com a frequência 0,132

Ordens diferentes de genes dão recombinantes duplos diferentes	
Ordens possíveis de genes	Cromátides de recombinantes duplos

FIG. 4.12 As três possíveis ordens de genes à esquerda produzem os seis produtos de um crossing duplo mostrado à direita. Apenas a primeira possibilidade é compatível com os dados do texto. Note que apenas as cromátides não-irmãs que tomam parte do crossing duplo são mostradas.

$\times 0,064 = 0,0084$ (0,84%). Na amostra de 1.448 moscas $0,0084 \times 1.448 = 12$ recombinantes duplos são esperados. Mas os dados mostram que apenas 8 foram de fato observados. Se essa deficiência de recombinantes duplos fosse consistentemente observada, isso nos mostraria que duas regiões não são independentes, e sugere que a distribuição de crossings favorece únicos à custa de duplos. Em outras palavras, há apenas um tipo de interferência: um crossing reduz a probabilidade de um crossing em uma região adjacente.

A interferência é quantificada primeiro calculando um termo chamado de **coeficiente de coincidência (c.o.c.)**, que é a proporção entre recombinantes duplos observados para esperados. A interferência (I) é definida como $1 - \text{c.o.c.}$. Assim,

$$I = 1 - \frac{\text{frequência observada ou número de recombinantes duplos}}{\text{frequência esperada ou número de recombinantes duplos}}$$

Em nosso exemplo,

$$I = 1 - \frac{8}{12} = \frac{4}{12} = \frac{1}{3}, \text{ ou } 33\%$$

Em algumas regiões, nunca foram observados recombinantes duplos. Nesses casos, $\text{c.o.c.} = 0$, e assim $I = 1$ e a interferência é completa. Os valores de interferência em qualquer ponto entre 0 e 1 são encontrados em regiões diferentes e em organismos diferentes.

Você deve ter indagado por que sempre usamos fêmeas heterozigotas para cruzamentos-teste em *Drosophila*. A explicação está em uma característica incomum dos machos de *Drosophila*. Quando, por exemplo, os machos $pr\ vg/pr^+ vg^+$ são cruzados com fêmeas $pr\ vg/pr\ vg$, apenas as proles $pr\ vg/pr^+ vg^+$ e $pr\ vg/pr\ vg$ são recuperadas. Esse resultado mostra que não há crossing em machos de *Drosophila*. Entretanto, essa ausência de crossing em um sexo é limitada a algumas espécies; não é o caso de machos de todas as espécies (ou do sexo heterogamético). Em outros organismos, há o crossing nos machos XY e nas fêmeas WZ. O motivo da ausência de crossing nos machos de *Drosophila* é que eles têm uma prófase I incomum, sem complexos sinaptonêmicos. Incidentalmente, há uma diferença de recombinação entre os sexos humanos também. As mulheres apresentam frequências mais altas de recombinantes para os mesmos loci que os homens.

Com o uso de uma reiteração das técnicas precedentes baseadas em recombinação, foram produzidos mapas de milhares de genes para os quais foram identificados fenótipos variantes (mutantes). Um simples exemplo ilustrativo de tomate é mostrado na Figura 4.13. Os cromossomos são mostrados como aparecem ao microscópio, junto com mapas cromossômicos baseados em análise de ligação de vários pares de alelos mostrados com seus fenótipos.

Uso de proporções como diagnóstico

A análise de proporções é um dos pilares da Genética. No texto até agora, encontramos muitas proporções diferentes

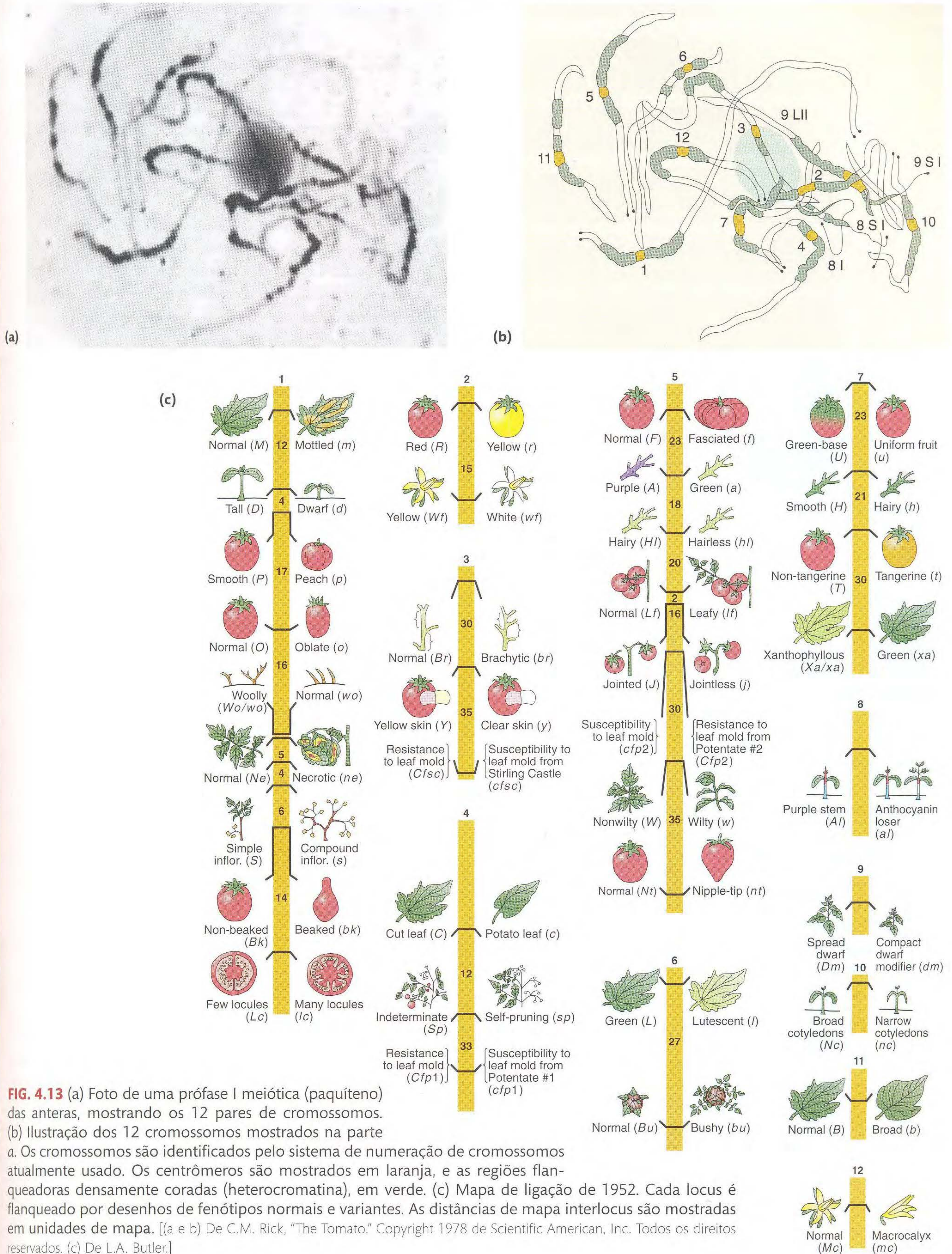
cujas derivações se espalham por vários capítulos. Como o reconhecimento das proporções e seu uso no diagnóstico do sistema genético em estudo são parte da genética cotidiana, vamos rever as principais proporções que vimos até agora. Elas são mostradas na Figura 4.14. Você pode ler as proporções pelas larguras relativas dos boxes coloridos em uma fileira. A Figura 4.14 lida com autofecundações e cruzamentos-teste de monóíbridos, diíbridos (com distribuição independente e ligação) e triíbridos (também com distribuição independente e ligação de todos os genes). Uma situação não representada é um triíbrido no qual apenas dois dos três genes estão ligados; como exercício, você pode gostar de deduzir o padrão geral que teria que ser incluído em tal diagrama dessa situação. Note que, com relação à ligação, os tamanhos das classes dependem das distâncias de mapa. Um geneticista deduz estados genéticos desconhecidos mais ou menos do seguinte modo: “uma proporção 9:3:3:1 me diz que essa proporção foi muito provavelmente produzida por um diíbrido autofecundado no qual os genes estão em cromossomos diferentes”.

► 4.3 Mapeamento com Marcadores Moleculares

A variação fenotípica baseada em alelos de genes é apenas uma parte da variação geral entre membros individuais de uma população. Como dito no Capítulo 2, um exame das seqüências de DNA de indivíduos diferentes de uma espécie revela uma variação considerável. Por exemplo, um indivíduo pode ter um par de bases GC em uma posição, digamos, 5.658 no DNA e outro indivíduo pode ter AT na posição 5.658. É possível, ainda, que o mesmo indivíduo tenha GC nessa posição em um membro de um par de cromossomos e AT na mesma posição no outro membro do par. Tal indivíduo é um heterozigoto molecular (“AT/GC”) para essa posição no DNA. Parte dessa variação afeta o fenótipo (por exemplo, mutações nulas dentro dos genes). Entretanto, uma proporção maior da variação de seqüência parece não estar relacionada ao fenótipo; em outras palavras, é uma variação *silenciosa*.

As variantes moleculares são tão comuns que muitos indivíduos são heterozigotos moleculares em vários loci diferentes no genoma. Esse fato é útil no mapeamento porque um heterozigoto molecular (“AT-GC”) pode ser mapeado como um heterozigoto fenotípico A/a . Um locus que constitui um heterozigoto molecular pode ser inserido em um mapa cromossômico analisando a frequência de recombinação exatamente do mesmo modo que é feita a inserção de um locus de alelos “fenotípicos” heterozigotos. Esse princípio é mantido muito embora a variação seja geralmente uma diferença silenciosa (talvez não em um gene). Esses loci de heterozigose molecular são chamados de **marcadores moleculares**. Agindo como importantes marcos em um mapa, os marcadores moleculares são úteis em orientar o pesquisador para encontrar o gene de interesse. Para compreender esse ponto, considere marcos reais: eles são de pouco interesse em si, mas são muito úteis em lhe dizer o quanto você está próximo de seu destino. Em um exemplo genético específico, suponha-

Mapa dos 12 cromossomos de tomate





PDF FEITO PELA SSECP - UFRN - NATAL/RN

Quase toda a diferença de 0,1% é baseada em diferenças de um único nucleotídeo. Como um exemplo, em um indivíduo, uma sequência localizada pode ser

....AAGGCTCAT....
....AACCGAGTA....

e, em outro, pode ser

....AAAGCTCAT....
....TTTCGAGTA....

Além disso, uma grande proporção dessas sequências localizadas são descobertas sendo polimórficas, significando que ambos os “alelos” são muito comuns na população. No geral, tais diferenças entre os indivíduos são chamadas de **polimorfismos de nucleotídeo único**, cuja abreviatura é **SNP** e a pronúncia, “**snip**”. Esses SNP têm grande potencial como marcos na análise de mapeamento. Em humanos, acredita-se que existam 3 milhões de SNP distribuídos mais ou menos aleatoriamente, com uma frequência de 1 em cada 300 a 1.000 bases, dando um conjunto útil de marcadores para mapeamento em escala fina.

Embora todos os SNP possam ser usados no mapeamento, alguns têm outros usos, dependendo de seu efeito a nível fenotípico. Existem várias categorias, entre as quais estão:

1. *SNP silenciosos dentro de genes.* Esses SNP não têm efeito aparente no fenótipo. Entretanto, são úteis como marcadores de alelos em particular. Por exemplo, se um morfo SNP está dentro de um alelo recessivo *a*, então esse alelo pode ser facilmente detectado usando o SNP.
2. *SNP nos quais um alelo de SNP causa um fenótipo mutante mostrando herança monogênica.* Esses SNP foram examinados na discussão de mutações no Capítulo 2. Em tais casos, em geral dentro da sequência codificante de proteínas, um “alelo” de SNP inativa ou altera a função gênica, resultando em algum tipo de fenótipo distinto. Em humanos, tal SNP pode estar associado a dois estados: um alelo de doença, tal como para PKU ou doença de Tay-Sachs, e o outro tipo selvagem. Aqui, novamente, o SNP pode ser usado como uma marcação; por exemplo, uma pessoa com o fenótipo *A/-* pode ser diagnosticada como heterozigota *A/a* detectando o morfo SNP associado.
3. *SNP em poligenes (QTL).* Exatamente do mesmo modo discutido para fenótipos monogênicos, esses SNP são úteis em descobrir poligenes que contribuem para um fenótipo apresentando variação contínua.
4. *SNP intergênicos.* Esses SNP não têm efeito fenotípico mensurável, mas ainda são úteis como marcos a caminho dos genes de interesse no modo geral descrito no começo desta seção (veja o Capítulo 2 e adiante, no presente capítulo).
5. *RFLP.* Já vimos os **polimorfismos de comprimento de fragmento de restrição (RFLP)**, que são SNP localizados no sítio-alvo de uma enzima de restrição. Em tais casos, existirão dois “alelos” de RFLP, ou morfos, um dos quais tem o alvo e o outro não. Os sítios de RFLP podem estar entre genes ou dentro desses. Como todos os SNP, eles são úteis como marcadores para o mapeamento de genes

de interesse. Embora relativamente complicados, eles têm a vantagem de não requerer seqüenciamento para sua detecção. Um exemplo de mapeamento de localização de um gene de interesse usando um RFLP é mostrado na Figura 4.15. Nesse exemplo, a frequência de recombinação é de 1 entre 8. Entretanto, cruzamentos mais similares precisarão ser feitos para obter uma amostra estatisticamente significativa. Na Figura 4.15, o SNP altera um sítio de restrição, mas qualquer SNP pode ser mapeado exatamente desse modo usando seqüenciamento mesmo que não altere um sítio de restrição.

Mapeamento usando haplótipos de SNP

Vimos que SNP *individuais* tais como RFLP podem ser usados como qualquer outro sítio heterozigoto para deduzir a ligação baseada na frequência de recombinantes. Entretanto, adicionalmente, as *combinações* de SNP podem ser usadas para deduzir ligação de um modo ligeiramente diferente, baseado na noção dos haplótipos. Um **haplótipo** é um segmento cromossômico definido por um arranjo específico de alelos SNP que ele leva. Assim, se usarmos primos (') para distinguir alelos alternativos de SNP, dois haplótipos definidos por cinco loci SNP podem ser:

.....1	2'	3'	4	5	e
.....1'	2	3	4'	5'	

Tais arranjos podem ser de qualquer tamanho, mas em geral estão entre 10 e 50 kb.

A nível populacional, os haplótipos são herdados por gerações como blocos. Com o passar do tempo, o crossing deverá romper os haplótipos, mas, em curto prazo, os haplótipos têm um tempo de vida finito à medida que aguardam a erosão pelo crossing. Durante esse tempo de vida finito, os alelos SNP de um haplótipo são ditos como tendo **desequilíbrio de ligação** (o estado de *equilíbrio* seria o obtido após crossings terem rompido o haplótipo). Devido ao **desequilíbrio de ligação**, os haplótipos podem ser usados para revelar a posição de mapa de um gene de interesse. Essa técnica é particularmente útil para mapear alelos de interesse médico no genoma humano, tais como alelos de doenças e alelos que predisõem a sensibilidade a drogas. Assim, um grande esforço está sendo feito em haplotipar populações humanas, e um mapa de haplótipo (“hapmap”) humano está sendo montado. Não foi surpresa descobrir pelo hapmap que, em humanos, os haplótipos são ilhas de baixa frequência de crossing separadas por pontos quentes de recombinação.

Para ver como os haplótipos podem ser usados para encontrar genes, considere o caso ilustrativo mostrado na Figura 4.16. No topo, vemos um haplótipo no qual, em alguma época no passado, ocorreu uma mutação em um gene de interesse, digamos, uma mutação de doença. À medida que o tempo passa, o crossing over introduz diferentes alelos de SNP, mas a *análise de associação* dos alelos de haplótipo na população atual mostra que os alelos originais 4 e 5 não foram recombinados e ainda estão em **desequilíbrio de ligação** com o alelo mutante. Assim, o locus do gene da doença pode ser localizado com bastante precisão porque os SNP humanos

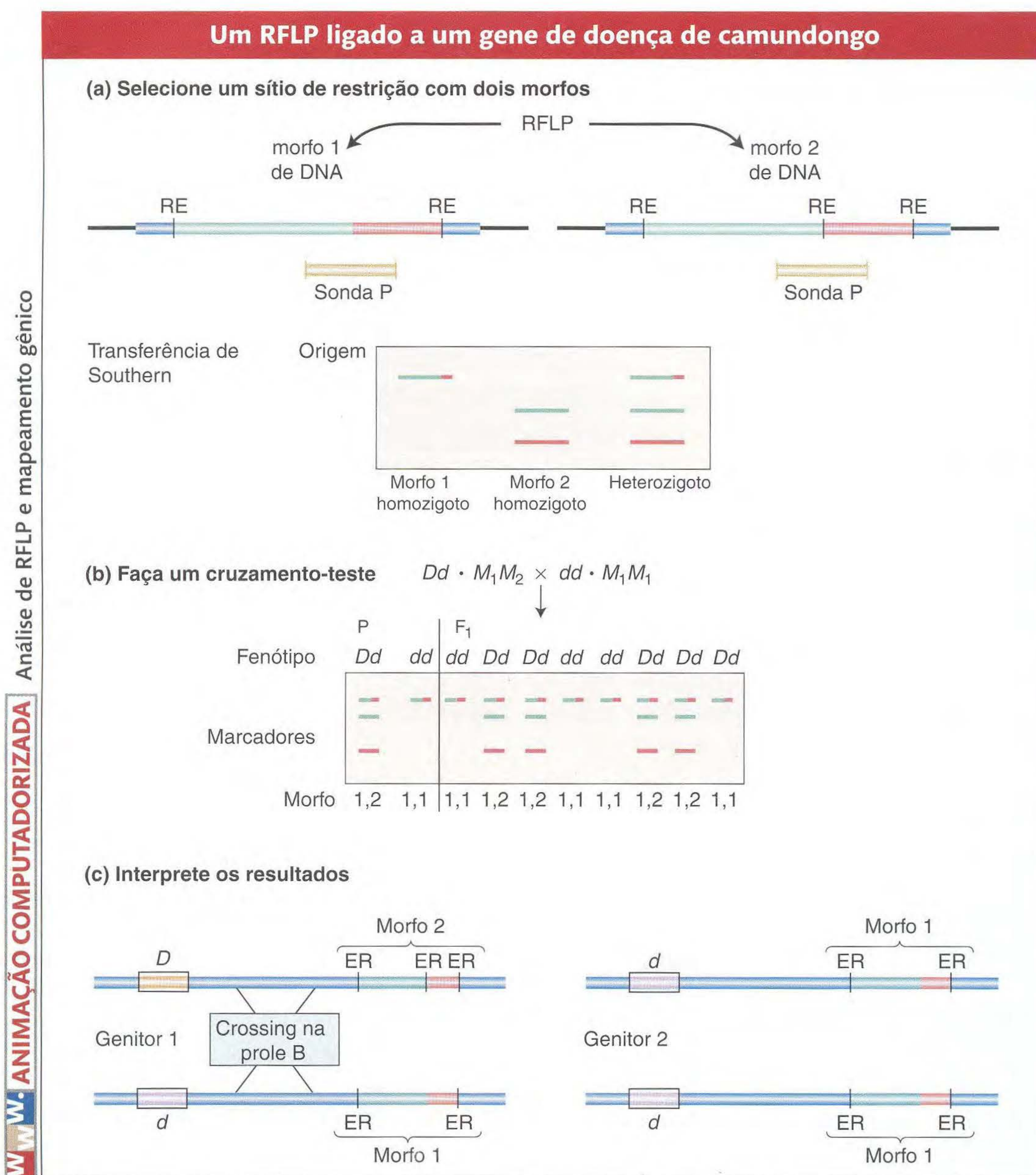


FIG. 4.15 A detecção e herança de um polimorfismo de comprimento de fragmento de restrição (RFLP). Uma sonda P detecta dois morfos de DNA quando o DNA é cortado por certa enzima de restrição (ER). O heredograma da doença dominante fenótipo D mostra ligação do locus D ao locus do RFLP; apenas a prole do indivíduo 8 é recombinante. Um mapa preciso iria requerer mais prole ou dados cumulativos de reproduções similares.

distam cerca de 1 kb. Tal análise de haplótipo foi usada para localizar, no cromossomo humano 4, um gene tendo um alelo que causa extrema sensibilidade à bronquiolite viral, um dos motivos mais comuns para hospitalização de crianças.

Você pode ver que o enfoque de desequilíbrio de ligação para mapear um gene, embora superficialmente diferente do mapeamento de frequência de recombinantes, é baseado no mesmo princípio, ou seja, para loci proximamente ligados, as combinações alélicas tendem a ser herdadas juntas.

SNP e análise de ligação de haplótipo são muito promissores para encontrar poligenes (QTL). Vimos que os poligenes fornecem um modelo genético para fenótipos que apresentam distribuição contínua tal como peso e altura. Os

poligenes também fornecem um modelo para a herança de condições humanas tais como fenda labial, pé torto, doença cardíaca, aterosclerose e hipertensão. Esses distúrbios apresentam uma herança complexa que pode ser explicada por um modelo poligênico no qual a condição não se expressa até que a dose de poligenes ultrapasse um *limiar*. SNP e a análise de haplótipos têm grande potencial para encontrar genes subjacentes tanto a variação contínua quanto a características com limiar.

Em organismos-modelo, os haplótipos podem ser usados para rastrear segmentos genômicos através de gerações e correlacionar esses segmentos à característica em questão. A idéia é ilustrada na Figura 4.17 com relação a linhagens

Uso de haplótipos para deduzir posição gênica

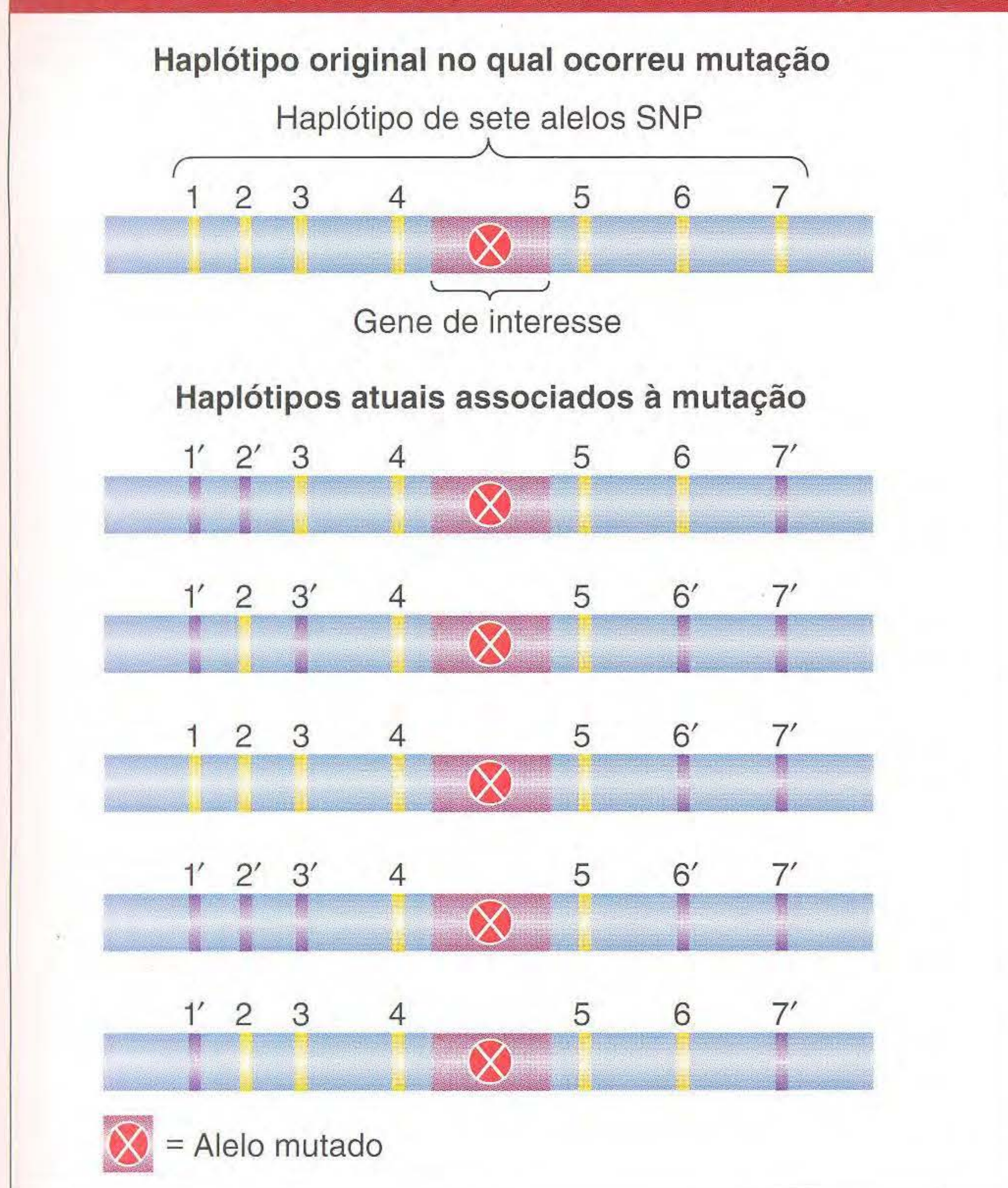


FIG. 4.16 O alelo mutante permanece em desequilíbrio de ligação apenas com SNP 4 e 5. Assim, o gene deve estar na vizinhança de 4 e 5.

puras de camundongos que diferem em alguma característica poligênica, tal como tendência a apresentar aterosclerose ou doença cardíaca. Pode ser feita uma correlação estatística de fenótipo com haplótipos que marca regiões cromossômicas específicas. Por exemplo, se o fenótipo da linhagem endocruzada do lado inferior direito da Figura 4.17 difere significativamente da linhagem parental B, então a região preta provavelmente contém um poligene contribuinte. É notável que a genética de camundongo tenha gerado modelos importantes para distúrbios de mamíferos, e muito do que aprendemos em camundongos pode ser aplicado aos humanos. Sem dúvida, isso também será verdade para os poligenes. Nos humanos, as similaridades de haplótipos em famílias ou populações podem ser usadas em análises estatísticas para detectar poligenes.

Polimorfismos de comprimento de seqüências simples

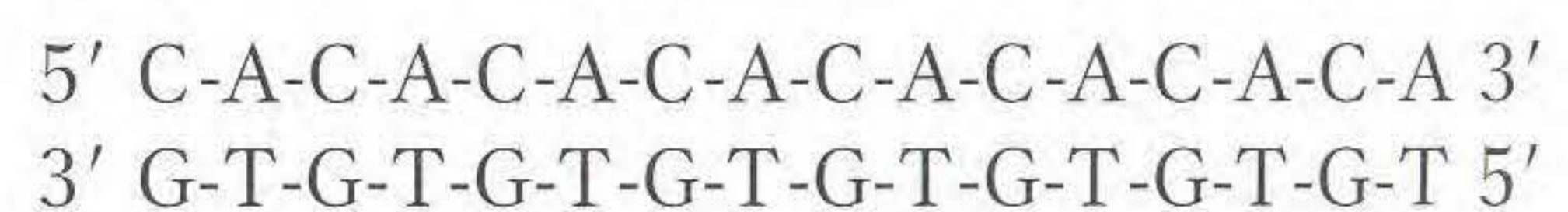
Uma das surpresas da análise genômica molecular é que a maioria dos genomas contém uma grande quantidade de DNA repetitivo. Além disso, existem muitos tipos de DNA repetitivo. Em uma ponta do espectro estão repetições múltiplas adjacentes de curtas seqüências simples de DNA. A origem dessas repetições não está clara, mas a característica que as torna úteis é que, em indivíduos diferentes, existem geralmente números diferentes de cópias. Assim, essas repetições são chamadas de **polimorfismos de comprimento de seqüência**

simples (SSLP). Elas também são chamadas, às vezes, **repetições em tandem de número variável** ou **VNTR**.

SSLP comumente têm alelos múltiplos; até 15 alelos foram encontrados para um locus de SSLP. Como consequência, às vezes 4 alelos (2 de cada genitor) podem ser rastreados em um heredograma. Dois tipos de SSLP são úteis no mapeamento e em outra análise do genoma: minissatélite e marcadores microssatélite. (A palavra *satélite* refere-se aqui à observação de que, quando o DNA genômico é isolado e fracionado com o uso de técnicas físicas, as seqüências repetitivas em geral formam uma fração que é fisicamente separada do resto; isto é, é uma fração satélite no sentido de que está separada da maior parte.)

Marcadores minissatélites Um **marcador minissatélite** é baseado na variação do número de repetições em tandem de uma unidade repetida de 15 a 100 nucleotídeos de tamanho. Em humanos, o tamanho total da unidade é de 1 a 5 kb. Os loci minissatélite tendo a mesma unidade repetida mas números de repetições diferentes são dispersos pelo genoma. Para definir um locus minissatélite, primeiro o DNA genômico total é cortado com uma enzima de restrição que não tem sítios-alvo dentro do arranjo, mas corta sítios flancadores; assim, o tamanho de cada fragmento corresponderá ao número de repetições. Se uma sonda que reconhece a seqüência repetitiva está disponível, então uma transferência de Southern revelará um grande número de fragmentos de tamanhos diferentes que são ligados pela sonda. De fato, esses padrões às vezes são chamados **fingerprints de DNA**. (Esses fingerprints são altamente individualistas e, assim, têm grande valor forense.) Para análise de recombinação, se os genitores diferem em determinada banda, então essa diferença torna-se um sítio heterozigoto que pode ser usado no mapeamento de um gene ligado de interesse. Um exemplo simples do uso de fingerprint de DNA para estabelecer a ligação é mostrado na Figura 4.18.

Marcadores microssatélites Um **marcador microssatélite** é baseado em números variáveis de repetições em tandem de uma seqüência mais simples, geralmente um dinucleotídeo. O tipo mais comum é uma repetição de CA e seu complemento GT, como no exemplo seguinte:



Os alelos microssatélites são revelados pela síntese de uma réplica da região repetitiva com o uso de uma técnica chamada de *reação em cadeia da polimerase (PCR)*. Essa técnica será explicada em mais detalhes no Capítulo 19, mas, essencialmente, ela usa um par de primers (“iniciadores de síntese”) de replicação de DNA que iniciam a síntese de DNA em um tubo de ensaio com o uso de uma seqüência genômica alvo como molde. Os primers são direcionais. No presente caso, dois primers são criados para “apontar” um para o outro de cada lado de um VNTR, levando o mecanismo de síntese de um lado para outro entre os primers e promovendo a síntese *in vitro* de um grande número de cópias da região.

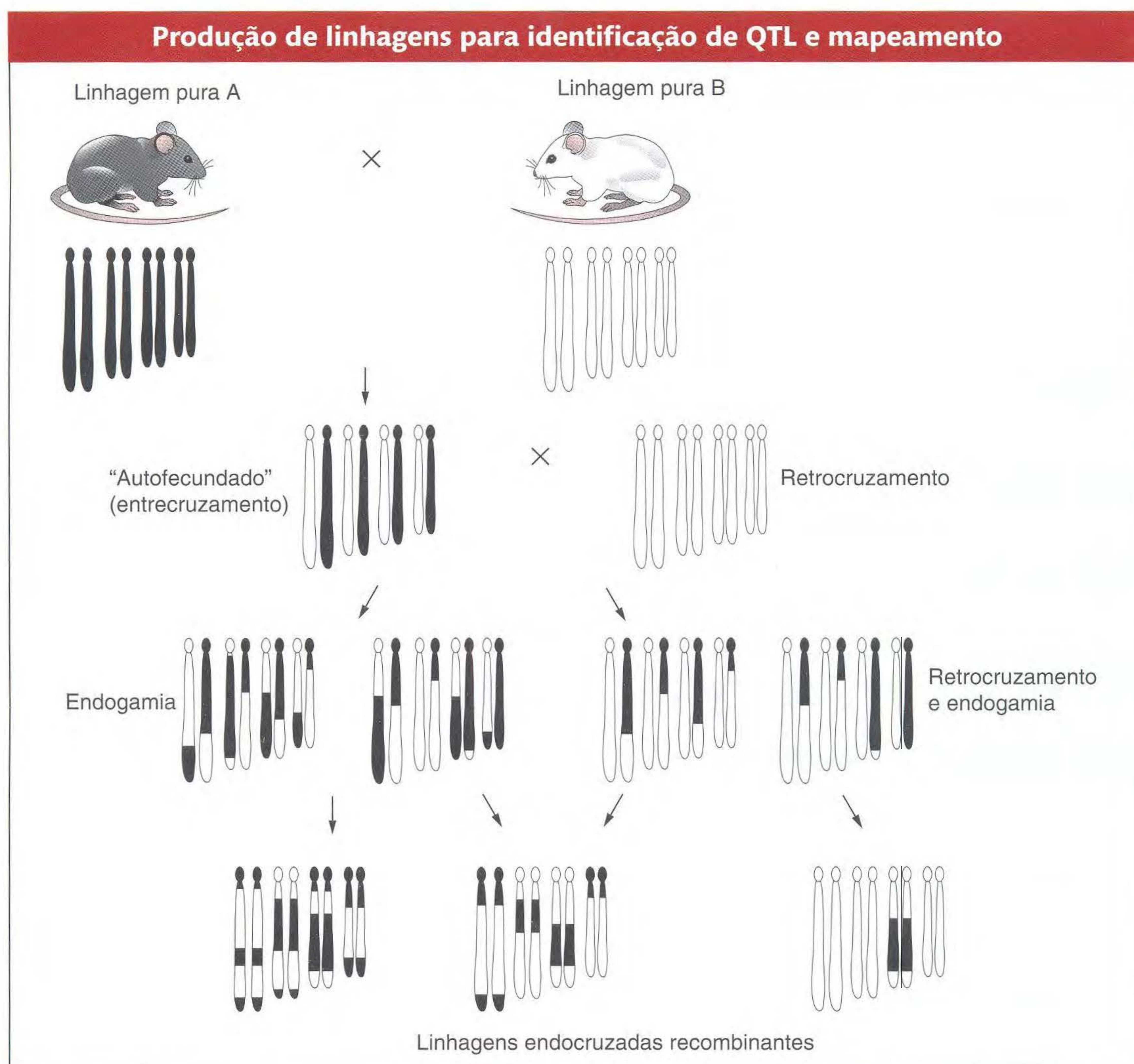
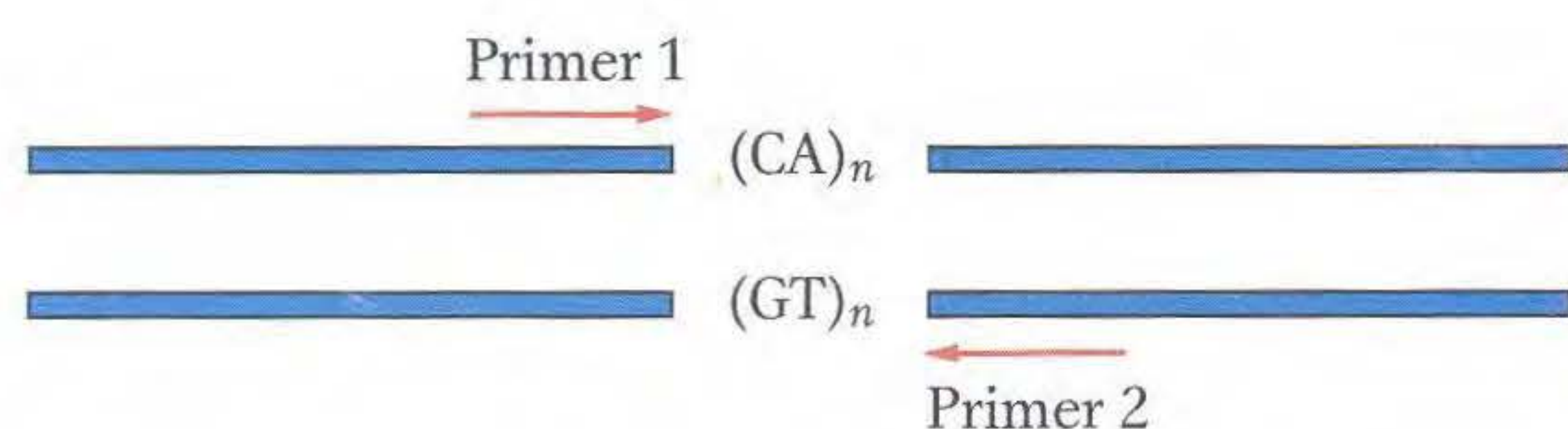


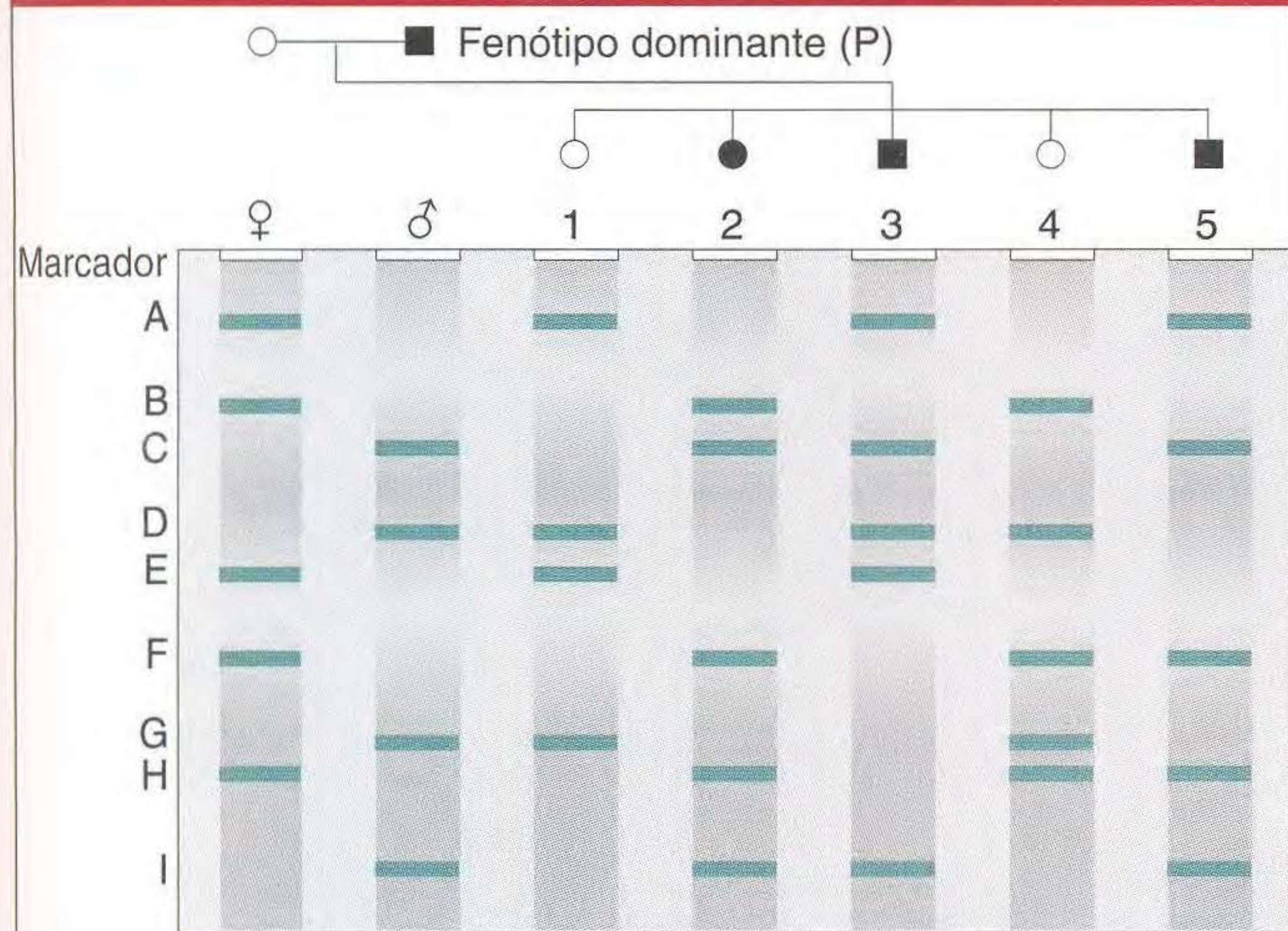
FIG. 4.17 Duas linhagens puras que diferem significativamente em alguma característica quantitativa são cruzadas, e, após a endogamia, são produzidas linhagens puras recombinantes. Uma comparação de fenótipos dessas linhagens leva à identificação de segmentos cromossomos específicos que, consistentemente, contribuem para a diferença parental. Essas regiões contêm supostos QTL. [Baseada em W.N. Frankel, "Taking Stock of Complex Trait Genetics in Mice." *Trends Genet.* 12, 1995, Fig. 1.]



Essa amostra amplificada de cópias idênticas pode ser detectada como uma banda em um gel de eletroforese. Qualquer variação de tamanho de microssatélite em indivíduos diferentes será revelada por bandas de tamanhos diferentes migrando para posições diferentes no gel. Assim, todas as variações de tamanho da repetição de microssatélite serão detectáveis. Uma alta proporção de tal análise de PCR revela vários "alelos" marcadores ou seqüências de repetições de tamanhos diferentes. Um exemplo de técnica de mapeamento de microssatélite é mostrado na Figura 4.19. Milhares de pares de primers de microssatélites podem ser feitos, os quais, do mesmo modo, detectam milhares de loci marcadores.

A Figura 4.20 contém alguns dados reais mostrando como os marcadores moleculares podem marcar um cromossomo humano. Você pode ver que o número de marcadores moleculares mapeados excede em muito o número de genes mapeados com fenótipos mutantes. Note que os SNP, devido à sua densidade ainda maior, não podem ser representados em um mapa de cromossomo inteiro tal como o da Figura 4.20, pois existiriam milhares deles. Um centimorgan (1 unidade de mapa) de DNA humano é um grande segmento, estimado como 1 megabase (1 Mb = 1 milhão de pares de bases, ou 1.000 kb). Assim, você pode ver a necessidade de marcadores moleculares proximalmente compactados para uma análise em escala fina que resolve distâncias menores. Note que o DNA equivalente a 1 unidade de mapa varia muito entre as espécies. Por exemplo, no parasita da malária *Plasmodium falciparum*, 1 unidade de mapa = 17 kb.

Algumas bandas de fingerprint de DNA mostram ligação



EXEMPLOS DE ANÁLISES

- F e H Sempre herdados juntos – ligados?
- A e B Na prole, sempre A ou B – “alélicos”?
- A e D Quatro combinações; A e D, A, D, ou nenhum – não-ligados?
- F, H e E Sempre ou F e H ou E – proximamente ligados em trans?
- Alelo P Possivelmente ligado a I e C.

FIG. 4.18 Cada banda contém uma região minissatélite flanqueada por cortes de restrição. As bandas I e C sempre aparecem nos fingerprints de indivíduos portadores do alelo dominante P, sugerindo ligação. (Nota: Essa reprodução não é um cruzamento-teste, apesar de revelar ligações.)

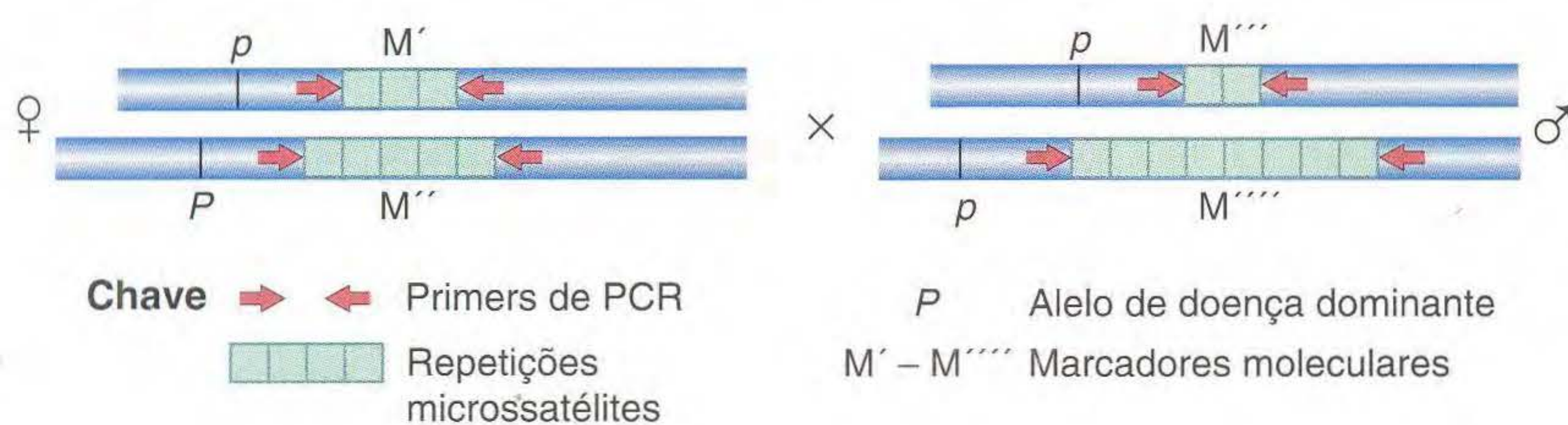
4.4 Mapeamento de Centrômero com Tétrades Lineares

Os centrômeros não são genes, mas sim regiões do DNA das quais a reprodução ordenada de organismos vivos depende absolutamente, e, portanto, são de grande interesse em genética. Na maioria dos eucariontes, a análise de recombinação não pode ser usada para mapear loci de centrômeros, pois eles não apresentam heterozigose que lhes permita serem usados como marcadores. Entretanto, nos fungos que produzem tétrades lineares (veja Capítulo 3), os centrômeros *podem* ser mapeados. Usaremos o fungo *Neurospora* como um exemplo. Lembre que, nos fungos tais como *Neurospora* (um haplóide), as divisões meióticas ocorrem ao longo do eixo longo do asco, e assim cada meiócito produz um arranjo linear de oito ascósporos, chamado de **óctade**. Esses oito ascósporos constituem os quatro produtos da meiose (uma tétrade) mais uma meiose pós-mitótica.

Em sua forma mais simples, o mapeamento do centrômero considera um locus gênico e pergunta a que distância esse locus está do centrômero. O método é baseado no fato de que um padrão diferente de alelos aparecerá em uma tétrade linear, ou óctade, que surge de uma meiose com um crossing entre um gene e seu centrômero. Considere um cruzamento entre dois indivíduos, cada um tendo um alelo diferente em um locus (digamos, $A \times a$). A lei de Mendel da segregação igual dita que, em uma óctade, sempre existirão quatro ascósporos com genótipo A e quatro com a. Mas como eles estão

Um locus de microsatélite pode mostrar ligação a um gene de doença

(a) Genótipos parentais



(b) Padrões de bandeamento de genitores e filhos

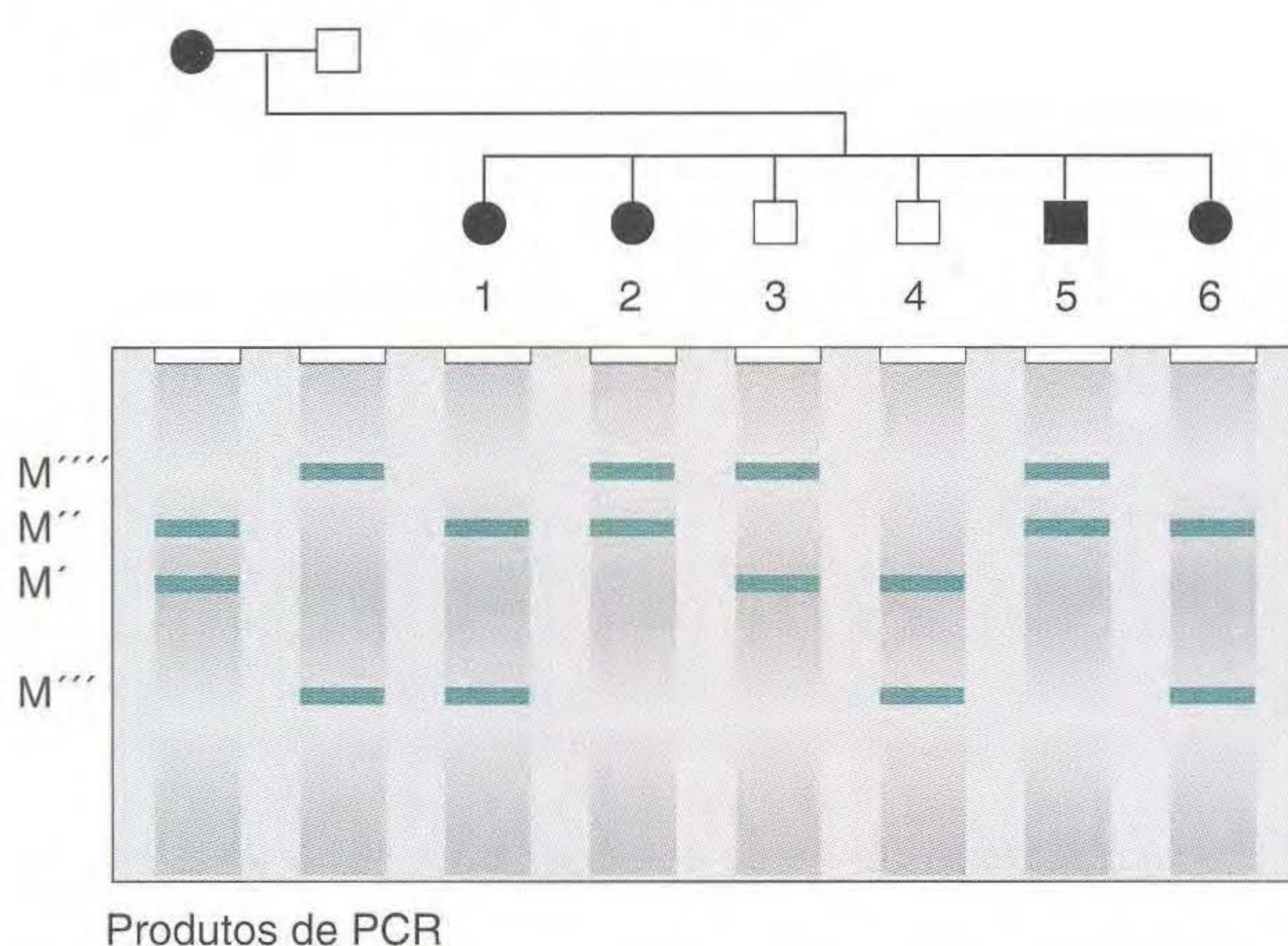
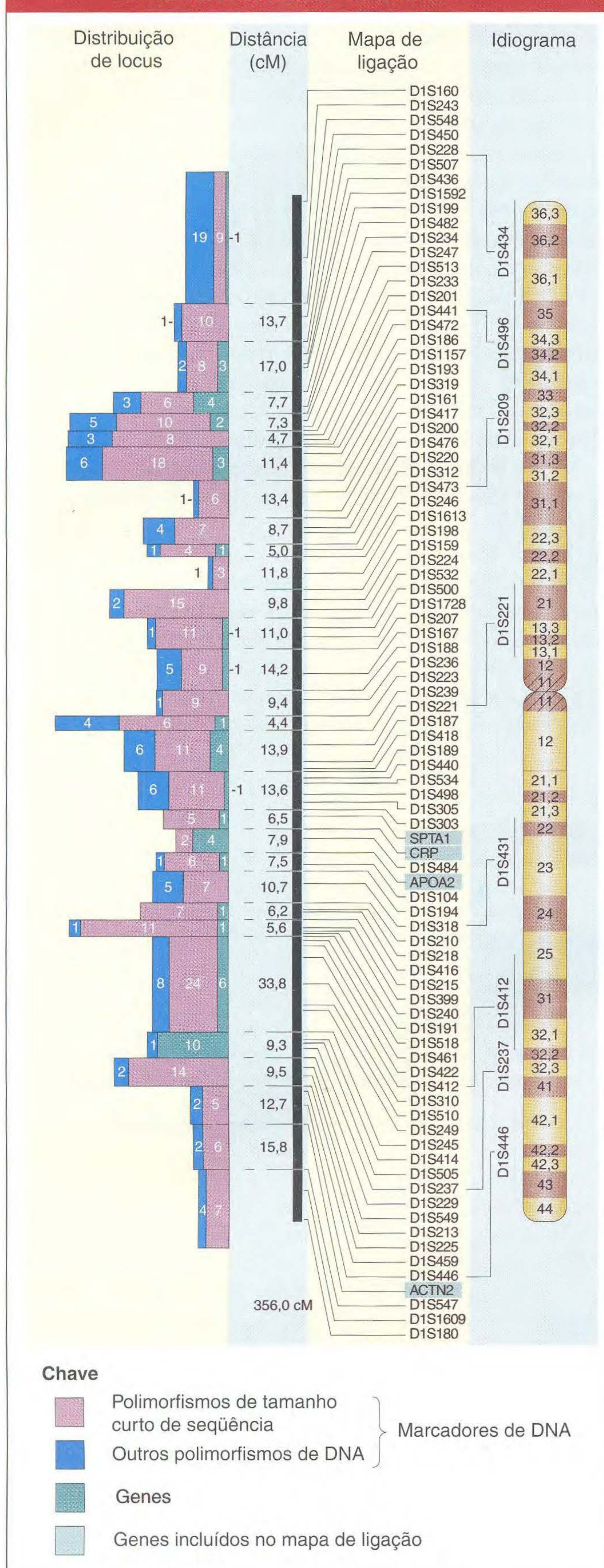


FIG. 4.19 Um padrão de bandeamento de PCR é mostrado para uma família com seis filhos, e esse padrão é interpretado no topo da ilustração com o uso de quatro “alelos” microsatélites de tamanhos diferentes, M' até M''' . Um desses marcadores (M'') provavelmente está ligado na configuração cis ao alelo P da doença. (**Nota:** Essa reprodução também não é um cruzamento-teste, embora seja informativa sobre ligação.)

Marcadores fenotípicos e moleculares mapeados no cromossomo humano 1



dispostos? Se não tiver havido crossing na região entre A/a e o centrômero, existirão dois blocos adjacentes de quatro ascoporos na óctade linear (veja Figura 3.10). Entretanto, se tiver havido um crossing nessa região, haverá um dentre quatro padrões diferentes na óctade, cada padrão mostrando *blocos de dois alelos adjacentes idênticos*. Alguns dados de um cruzamento real de $A \times a$ são mostrados no quadro seguinte.

Óctades					
A	a	A	a	A	a
A	a	A	a	A	a
A	a	a	A	a	A
A	a	a	A	a	A
a	A	A	a	a	A
a	A	A	a	a	A
a	A	a	A	A	a
a	A	a	A	A	a
126	132	9	11	10	12
Total = 300					

As primeiras duas colunas da esquerda são de meioses *sem crossing* na região entre o locus A e o centrômero. Os padrões para essas colunas são chamados de **padrões de segregação de primeira divisão (padrões M_I)** porque os dois alelos diferentes segregam para os dois núcleos-filhos na primeira divisão de meiose. As outras quatro colunas são todas de meiócitos com um crossing. Esses padrões são chamados de **padrões de segregação de segunda divisão (M_{II})**, pois, como resultado do crossing na região centrômero para o locus, os alelos A e a ainda estão juntos nos núcleos ao final da primeira divisão de meiose (Figura 4.21). Não houve segregação de primeira divisão. Entretanto, a segunda divisão de meiose segregava os alelos A e a para núcleos separados. A Figura 4.21 mostra como um desses padrões M_{II} é produzido. Os outros padrões são similarmente produzidos. A diferença é que as cromátides movem-se em direções diferentes na segunda divisão (Figura 4.22).

Você pode ver que a frequência de óctades com um padrão M_{II} deve ser proporcional ao tamanho do centrômero-região

FIG. 4.20 O diagrama mostra a distribuição de todas as diferenças genéticas que foram mapeadas no cromossomo 1 na época em que esse diagrama foi desenhado. Alguns marcadores são genes de fenótipos conhecidos (seus números são sombreados em verde), mas a maioria é de marcadores polimórficos de DNA (os números sombreados em rosa e azul representam duas classes diferentes de marcadores moleculares). Um mapa de ligação apresentando um conjunto bem espaçado desses marcadores, com base nas análises de frequência de recombinantes do tipo descrito neste capítulo, está no centro da ilustração. As distâncias de mapa são mostradas em centimorgans (cM). Com uma distância total de 356 cM, o cromossomo 1 é o maior cromossomo humano. Alguns marcadores também foram localizados no cromossomo 1 do mapa citogenético (mapa à direita, chamado de *idiograma*), usando técnicas descritas mais adiante neste capítulo. Havendo marcadores comuns em mapas genéticos diferentes, permite as localizações de outros genes e marcadores moleculares a serem estimados em cada mapa. A maioria dos marcadores mostrados no mapa de ligação são moleculares, mas vários genes (destacados em verde-claro) também estão incluídos. [B.R. Jasney et al., *Science*, September 30, 1994.]

Padrão de segregação de segunda divisão em uma óctade de fungo

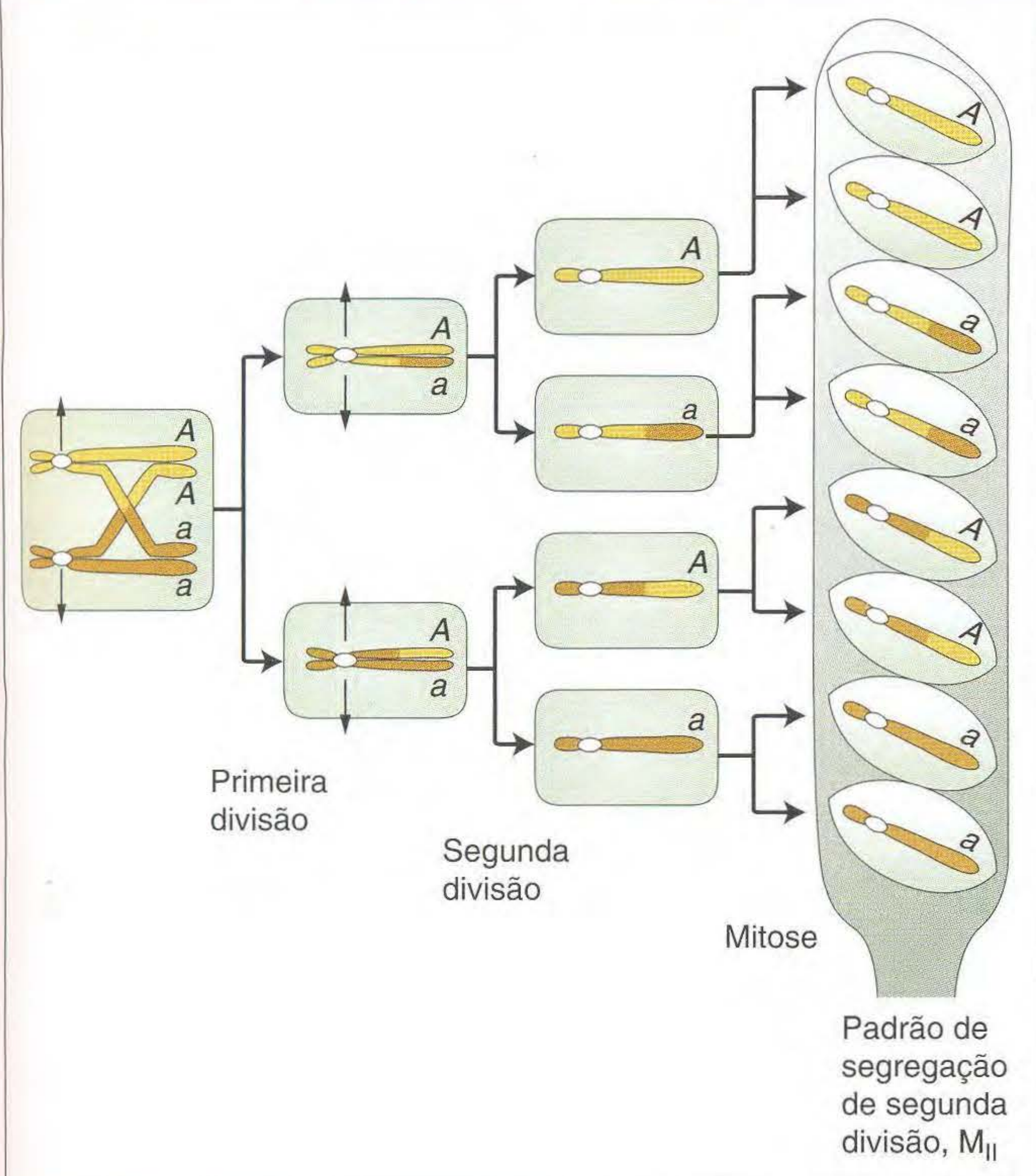


FIG. 4.21 A e a segregam-se em núcleos separados na segunda divisão de meiose quando há crossing entre o centrômero e o locus A.

A/a e pode ser usado como uma medida do tamanho dessa região. Em nosso exemplo, a frequência de M_{II} é $42/300 = 14\%$. Essa percentagem significa que o locus de tipo reprodutivo está 14 unidades de mapa do centrômero? A resposta é não, mas esse valor pode ser usado para calcular o número de unidades de mapa. O valor de 14% é uma percentagem da *meiose*, que não é o modo pelo qual as unidades de mapa são definidas. As unidades de mapa são definidas como a percentagem de *cromátides* recombinantes vindas da meiose. Como um crossing em qualquer meiose resulta apenas em 50% de cromátides recombinantes (quatro de oito; veja Figura 4.21), devemos dividir os 14% por 2 para converter a frequência M_{II} (uma frequência de *meiose*) em unidades de mapa (uma

frequência de *cromátides* recombinantes). Assim, essa região deve ter 7 unidades de mapa de tamanho, medida que pode ser introduzida no mapa desse cromossomo.

4.5 Uso do Teste do Qui-quadrado para Testar Análise de Ligação

Na análise de ligação, sempre surge a seguinte pergunta: Estes dois genes estão ligados? Às vezes a resposta é óbvia porque a frequência de recombinantes é substancialmente menor que 50%, e às vezes não. Mas, em ambas as situações, a aplicação de um teste estatístico objetivo que pode apoiar ou não a instituição é útil. O teste do χ^2 , que foi visto no Capítulo 3, fornece um modo útil para decidir se dois genes estão ligados. Como o teste do χ^2 é aplicado na ligação? Como já discutido neste capítulo, podemos deduzir que dois genes estão ligados no mesmo cromossomo se a FR é menor que 50%. Mas, e quanto a valores próximos de 50%?

Vamos testar um conjunto específico de dados para ligação usando a análise de χ^2 . Suponha que cruzamos genitores puros com os genótipos $A/A \cdot B/B$ e $a/a \cdot b/b$ e obtivemos um diíbrido $A/a \cdot B/b$, o qual usaremos em um cruzamento-teste com $a/a \cdot b/b$. Uma prole de 500 foi classificada do seguinte modo (escrita como gametas do diíbrido):

142	$A \cdot B$	parental
133	$a \cdot b$	parental
113	$A \cdot b$	recombinante
112	$a \cdot B$	recombinante

Total: 500

A partir desses dados, a frequência de recombinantes é $225/500 = 45\%$. Face a isso, esse resultado parece devido à ligação porque a RF é menor que os 50% esperados da distribuição independente. Entretanto, é possível que os dois genes não estejam ligados e as classes recombinantes sejam menores que os 50% apenas por acaso. Assim, precisamos fazer um teste de χ^2 para calcular a probabilidade desse resultado devido ao acaso.

Como é usual para o teste do χ^2 , a primeira etapa é calcular as expectativas, E , para cada classe. Ao formular E , imediatamente entramos em um problema: não temos uma distância de ligação precisa para usar na formulação de nossas expectativas se a hipótese “os genes estão ligados” é verdadeira.

Quatro ligações diferentes ao fuso produzem quatro padrões de segregação de segunda divisão

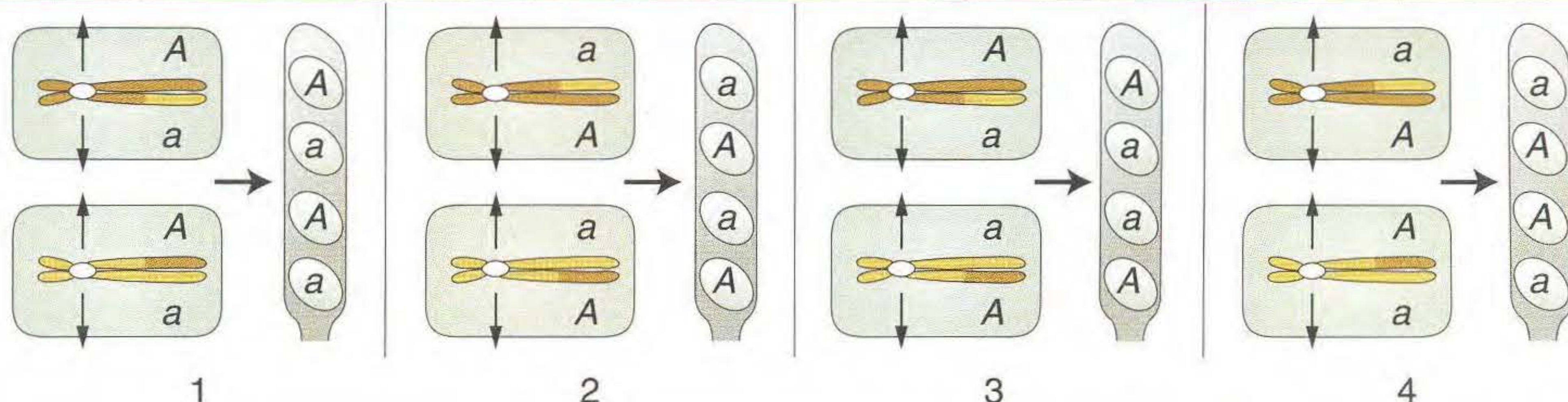
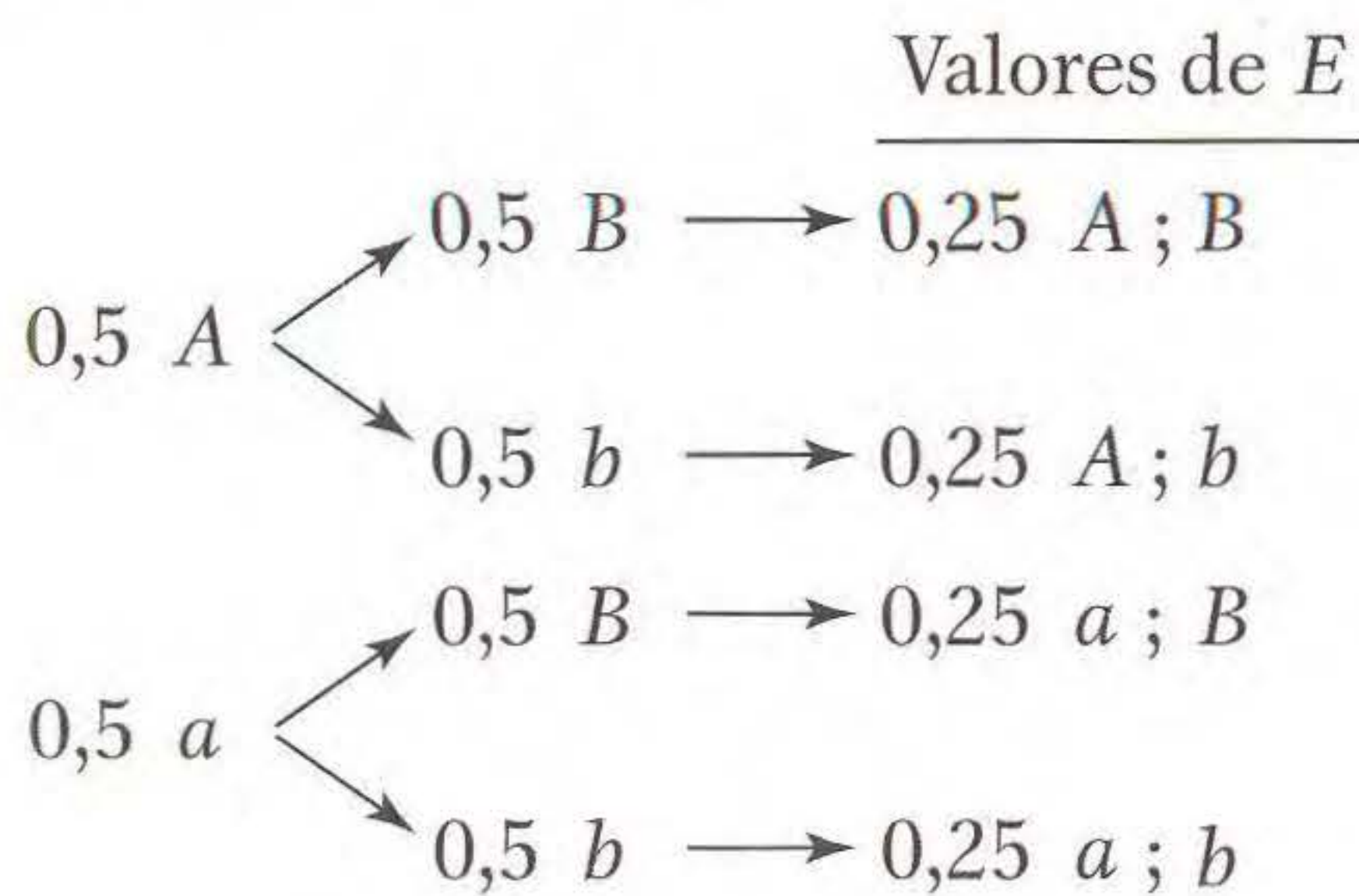


FIG. 4.22 Na segunda divisão de meiose, os centrômeros ligam-se aleatoriamente ao fuso, produzindo os quatro arranjos mostrados. Os quatro arranjos são igualmente frequentes.

Os genes distam 1 u.m.? 10 u.m.? 45 u.m.? Portanto, não podemos testar a ligação diretamente. Entretanto, a hipótese que podemos usar para fazer uma previsão exata é a de distribuição independente, ou seja, a ausência de ligação. Se os resultados observados nos fizerem rejeitar a hipótese de não-ligação, então nós podemos deduzir ligação. Esse tipo de hipótese, chamada de **hipótese nula**, é geralmente útil na análise do χ^2 porque fornece uma previsão experimental exata que pode ser testada.

Como podemos calcular os valores *E* gaméticos para a hipótese de que não há ligação? Um meio pode ser fazer uma simples previsão com base na Primeira e Segunda Leis de Mendel, do seguinte modo:



Assim, podemos assegurar que, se os pares de alelos do diíbrido estão se distribuindo independentemente, deve haver uma proporção de tipos gaméticos de 1:1:1:1. Portanto, usando 1/4 de 500, ou 125, como a proporção esperada de cada classe gamética, parece razoável. Entretanto, note que a proporção 1:1:1:1 só é esperada se todos os genótipos forem igualmente viáveis. Geralmente, os genótipos *não são* igualmente viáveis, porque os indivíduos que levam alguns alelos não sobrevivem até a vida adulta. Portanto, em vez das proporções de alelos de 0,5:0,5, usadas antes, podemos ver, por exemplo, proporções de 0,6 *A*:0,4 *a* ou 0,45 *B*:0,55 *b*. Devemos usar essas proporções em nossas previsões de independência.

Vamos arranjar as classes genotípicas observadas em uma grade para revelar mais claramente as proporções alélicas.

VALORES OBSERVADOS

		Segregação de <i>A</i> e <i>a</i>		
		<i>A</i>	<i>a</i>	Total
Segregação de <i>B</i> e <i>b</i>	<i>B</i>	142	112	254
	<i>b</i>	113	133	246
Total		255	245	500

Vemos que as proporções de alelos são $\frac{255}{500}$ para *A*, $\frac{245}{500}$ para *a*, $\frac{254}{500}$ para *B* e $\frac{246}{500}$ para *b*. Agora, calculamos os valores esperados sob distribuição independente simplesmente multiplicando essas proporções alélicas. Por exemplo, para encontrarmos o número esperado de genótipos *A B* na amostra se duas proporções são combinadas aleatoriamente, nós simplesmente multiplicamos os seguintes termos:

$$\text{Valor esperado, ou } E, \text{ para } A B = (255/500) \times (254/500) \times 500 = 129,54$$

Com o uso desse enfoque, toda a grade de valores *E* pode ser completada, do seguinte modo:

VALORES ESPERADOS

		Segregação de <i>A</i> e <i>a</i>		
		<i>A</i>	<i>a</i>	Total
Segregação de <i>B</i> e <i>b</i>	<i>B</i>	129,54	124,46	254
	<i>b</i>	125,46	120,56	246
Total		255	245	500

O valor do χ^2 é calculado do seguinte modo (no qual *O* é o valor observado):

Genótipo	<i>O</i>	<i>E</i>	(<i>O</i> - <i>E</i>) ² / <i>E</i>
<i>A B</i>	142	129,54	1,19
<i>a b</i>	133	120,56	1,29
<i>A b</i>	113	125,46	1,24
<i>a B</i>	112	124,46	1,25
Total (que é igual ao valor do χ^2) = 4,97			

O valor obtido de χ^2 (4,97) é usado para encontrar o valor da probabilidade correspondente, *p*, usando a tabela do χ^2 (veja Capítulo 3, Quadro 3.1). Primeiro, precisamos estabelecer os graus de liberdade em ordem para escolher a fileira correta no quadro. Geralmente, em um teste estatístico, o número de graus de liberdade é o número de valores não-dependentes. Trabalhando com o seguinte “experimento imaginário”, mostraremos o que essa afirmativa significa na presente aplicação. Em uma grade de dados 2 × 2 (do tipo usado antes), como os totais de colunas e filas são dados de resultados experimentais, a especificação de qualquer valor dentro da grade automaticamente dita os outros três valores. Assim, só há um valor não-dependente e, portanto, apenas um grau de liberdade. Uma regra geral para grades maiores é que o número de graus de liberdade é igual ao número de classes representadas nas fileiras menos um multiplicado pelo número de classes representadas nas colunas menos um. Aplicando essa regra no exemplo atual, temos

$$df = (2 - 1) \times (2 - 1) = 1$$

Portanto, usando o Quadro 3.1 (Capítulo 3), procuramos ao longo da fileira correspondente a um grau de liberdade até localizar nosso valor de χ^2 de 4,97. Nem todos os valores de χ^2 são mostrados no Quadro 3.1, mas 4,97 está perto do valor de 5,021. Logo, o valor de probabilidade correspondente está muito próximo de 0,025, ou 2,5%. Esse valor de *p* é o valor da probabilidade que procuramos, o de obter um desvio esperado desse tamanho ou maior. Como essa probabilidade é menor que 5%, a hipótese de distribuição independente deve ser rejeitada. Assim, tendo rejeitado a hipótese de não-ligação, ficamos com a dedução de que de fato os loci estão provavelmente ligados.

► 4.6 Uso de Valores Lod para Avaliar a Ligação em Heredogramas Humanos

Os humanos têm milhares de fenótipos herdados autossomica-mente, e o mapeamento dos loci dos genes que causam esses fenótipos pode parecer uma simples questão de usar recombi-nação, como mostrado neste capítulo. Entretanto, o progresso no mapeamento desses loci foi inicialmente lento por vários motivos. Primeiro, os cruzamentos controlados não podem ser feitos com seres humanos, e os geneticistas tiveram que tentar calcular frequências de recombinantes em diíbridos ocasionais que foram produzidos ao acaso em reproduções humanas. As reproduções que eram equivalentes a cruzamentos-teste eram extremamente raras. Segundo, os seres humanos em geral pro-duzem apenas um pequeno número de prole, dificultando a obtenção de dados suficientes para calcular distâncias de mapa estatisticamente confiáveis.

Para obter valores de FR confiáveis, são necessárias gran-des-amostras. Entretanto, mesmo quando o número de prole de qualquer reprodução individual é pequeno, pode ser feita uma estimativa mais confiável combinando os resultados de muitas reproduções idênticas. O modo padrão de fazer isso é calcular os **valores Lod** (*Lod* quer dizer “logaritmo das pro-babilidades [*odds*])). O método simplesmente calcula duas probabilidades diferentes de obter um conjunto específico de recombinantes observado em uma família. A primeira probabilidade é calculada com a suposição de distribuição independente, e a segunda probabilidade é calculada com a suposição de um grau específico de ligação. Então a propor-ção das probabilidades é calculada, e o logaritmo desse núme-ro é obtido, que é o valor Lod. O procedimento é repetido para uma gama de vários graus diferentes de ligação. O grau de ligação que produz o maior valor Lod é o mais provável. Vejamos um simples exemplo de como o cálculo funciona.

Suponha que temos uma família para um cruzamento-tes-te diíbrido. Suponha também que, para o indivíduo diíbrido, podemos deduzir os gametas de entrada e, assim, avaliar os gametas individuais para recombinação. O diíbrido é hete-rozigoto para um alelo de doença dominante (*D/d*) e para um marcador molecular (*M1/M2*). Suponha tratar-se de um homem e que os gametas que se uniram para formá-lo foram *D · M1* e *d · M2*. Sua esposa é *d/d · M2/M2*. O heredograma na Figura 4.23 mostra seus seis filhos, categorizados como parentais ou recombinantes com relação ao gameta contri-buinte do pai. Dos seis, existem dois recombinantes, que dariam uma FR de 33%. Tal FR, que é bem menor que 50%, sugere que o gene da doença e o marcador estão ligados, com uma separação entre 30 e 40 unidades de mapa. Entretanto, o gene e o marcador possivelmente não estão ligados; ou melhor, distribuem-se independentemente, mas o tamanho pequeno da amostra gerou um resultado enganador. Vamos calcular a probabilidade desse resultado sob várias hipóteses. O quadro seguinte mostra as proporções esperadas de genó-tipos parentais (P) e recombinantes (R) sob três valores de FR e sob distribuição independente:

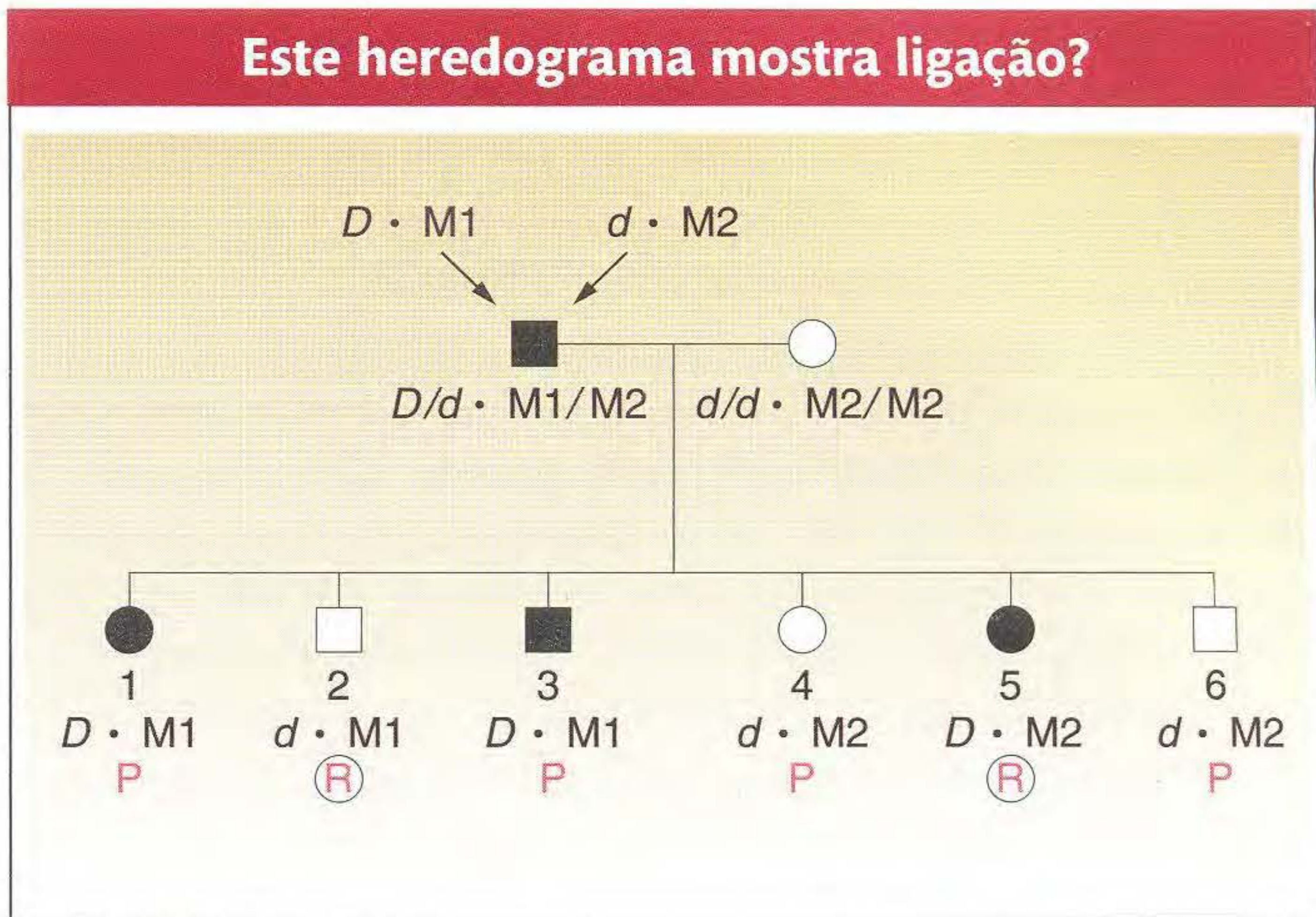


FIG. 4.23 O heredograma é efetivamente um cruzamento-teste. *D/d* são alelos de um gene de doença; M1 e M2 são “alelos” moleculares, tais como duas formas de um RFLP. P, parental (não-recombinante); R, recombinante.

	FR			
	0,5	0,4	0,3	0,2
P	0,25	0,3	0,35	0,4
P	0,25	0,3	0,35	0,4
R	0,25	0,2	0,15	0,1
R	0,25	0,2	0,15	0,1

A probabilidade de obter os resultados sob distribuição independente (FR de 50%) será igual a

$0,25 \times 0,25 \times 0,25 \times 0,25 \times 0,25 \times 0,25 \times B = 0,00024 \times B$
onde B = número de possíveis ordens de nascimento para 4 parentais e dois indivíduos recombinantes.

Para uma FR de 0,2 (20%), a probabilidade é

$$0,4 \times 0,1 \times 0,4 \times 0,4 \times 0,1 \times 0,4 \times B = 0,00026 \times B$$

A proporção dos dois é $0,00026/0,00024 = 1,08$ (note que os B foram eliminados). Assim, com base nos dados dessa família, a hipótese de uma FR de 0,2 é 1,08 vez tão provável quanto a hipótese de distribuição independente. Finalmente, tomamos o logaritmo da proporção para obter o valor Lod. Algumas outras proporções e seus valores Lod são mostrados no seguinte quadro:

	FR			
	0,5	0,4	0,3	0,2
Probabilidade	0,00024	0,00032	0,00034	0,00026
Proporção	1,0	1,33	1,41	1,08
Lod	0	0,12	0,15	0,03

Não surpreendentemente, esses números mostram nos-sa suspeita original de que a FR é de cerca de 30 a 40% porque essas hipóteses geram os maiores valores Lod. Mas esse resultado apenas não dá um suporte convincente para qualquer modelo de ligação. Ao contrário, à medida que os dados de mais reproduções desse genótipo se tornam dispo-

níveis, suas proporções de possibilidades podem ser combinadas para dar um teste melhor de qualquer das hipóteses de ligação. Por exemplo, suponha que encontramos que, em cinco reproduções independentes diferentes, as proporções das possibilidades para uma FR de 20% são 1,08, 1,20, 1,05, 1,32 e 1,15. As possibilidades gerais agora seriam o produto dessas possibilidades individuais, essencialmente o produto de todos os numeradores dividido pelo produto de todos os denominadores, um cálculo análogo à regra do produto. O produto é 2,07, mostrando que o modelo de 20% de ligação é cerca de duas vezes mais provável que a independência.

Como os logaritmos são expoentes, na prática somar os valores Lod de reproduções individuais é mais conveniente. Desse modo, pesquisadores individuais podem facilmente somar seus valores aos acumulados por outros. Convencionalmente, um valor Lod cumulativo de pelo menos 3 é considerado um apoio convincente para um valor FR específico. Note que, como os logaritmos são expoentes na base 10, um valor Lod de 3 representa um valor de FR que é 1.000 vezes (isto é, 10^3) mais provável que a hipótese de não haver ligação, e, portanto, esse teste é bem rigoroso. O valor Lod de 20% de recombinação dos dados combinados no parágrafo precedente é 0,3148, ainda uma longa distância abaixo do valor 3.

► 4.7 Cômputo de Crossings Múltiplos Não-vistos

Na discussão do cruzamento-teste de três pontos, algumas cromátides parentais (não-recombinantes) resultam de crossings *duplos*. Esses crossings inicialmente não podiam ser contados na frequência de recombinantes, desviando os resultados. Essa situação levou a uma noção preocupante de que *todas* as distâncias de mapa baseadas em frequência de recombinantes podem ser subestimativas de distâncias físicas porque os crossings múltiplos não-detectados podem ter ocorrido, alguns de cujos produtos não seriam recombinantes. Vários enfoques matemáticos criativos foram criados para contornar o problema de crossing múltiplo. Veremos dois métodos. Primeiro, examinaremos um método originalmente desenvolvido por J.B.S. Haldane nos primeiros anos da Genética.

Função de Mapa

O enfoque desenvolvido por Haldane foi criar uma **função de mapa**, uma fórmula que relaciona um valor observado de frequência recombinante com uma distância de mapa corrigida para múltiplos crossings. O enfoque funciona relacionando FR ao número médio de crossings, m , que deve ter ocorrido nesse segmento cromossômico por meiose e então deduzindo que distância de mapa esse valor m deve ter produzido.

Para encontrarmos a relação de FR com m , devemos primeiro pensar nos resultados das várias possibilidades de crossing. Em qualquer região cromossômica, devemos esperar meioses com zero, um, dois, três, quatro ou mais crossings. Surpreendentemente, a única classe que é realmente crucial

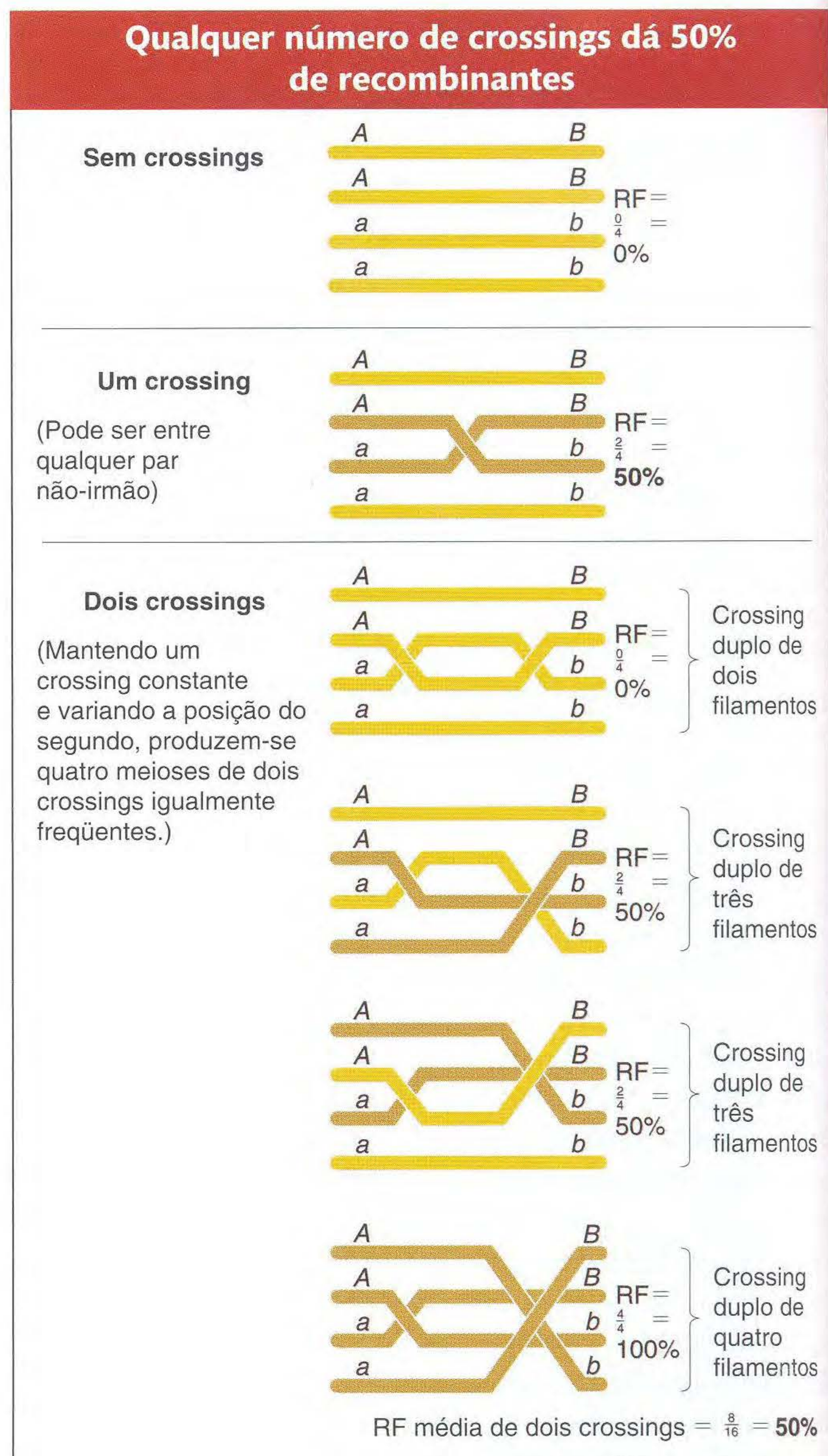


FIG. 4.24 Demonstração de que a FR média é 50% das meioses nas quais o número de crossings não é zero. As cromátides recombinantes são marrons. Crossings duplos com dois filamentos produzem todos os tipos parentais; assim, todas as cromátides são laranja. Note que todos os crossings são entre cromátides não-irmãs. Tente você mesmo a classe de crossing triplo.

é a classe zero. Para saber o porquê, considere o seguinte. É curioso, mas não um fato intuitivo, que *qualquer número* de crossings produz uma frequência de 50% de recombinantes *dentro dessas meioses*. A Figura 4.24 prova essa afirmativa para crossings únicos e duplos como exemplos, mas é verdade para qualquer número de crossings. Assim, o determinante verdadeiro da FR são os tamanhos relativos das classes sem crossing (a classe zero) comparada com as classes com qualquer número não-zero de crossings.

Agora a tarefa é calcular o tamanho da classe zero. A ocorrência de crossings em uma região cromossômica específica é bem descrita por uma distribuição estatística chamada de **distribuição de Poisson**. A fórmula de Poisson em geral

descreve a distribuição dos “sucessos” nas amostras quando a probabilidade média de sucessos é baixa. Um exemplo ilustrativo é mergulhar uma rede em um lago com peixes. A maioria das imersões não produzirá nenhum peixe, uma pequena proporção produzirá um peixe e mesmo uma proporção menor produzirá dois, e assim por diante. Esse análogo pode ser diretamente aplicado a uma região cromossômica, que terá 0, 1, 2, e assim por diante, “sucessos” de crossing em meioses diferentes. A fórmula de Poisson dada aqui nos dirá a proporção das classes com números diferentes de crossings.

$$f_i = (e^{-m} m^i) / i!$$

Os termos na fórmula têm os seguintes significados:

e = a base dos logaritmos naturais (aproximadamente 2,7)

m = o número médio de sucessos em um tamanho definido de amostra

i = o número real de sucessos em uma amostra desse tamanho

f_i = a frequência de amostras com i sucessos nelas

$!$ = o símbolo de fatorial (por exemplo, $5! = 5 \times 4 \times 3 \times 2 \times 1$)

A distribuição de Poisson nos diz que a frequência de i = classe 0 (a chave) é

$$e^{-m} \frac{m^0}{0!}$$

Como m^0 e $0!$ são ambos iguais a 1, a fórmula reduz-se a e^{-m} .

Agora podemos escrever uma função que relaciona FR a m . A frequência da classe com qualquer número não-zero de crossings será $1 - e^{-m}$, e, nessas meioses, 50% (1/2) dos produtos serão recombinantes; logo,

$$FR = \frac{1}{2} (1 - e^{-m})$$

e essa fórmula é a função de mapa que estávamos procurando.

Vejamos um exemplo no qual FR seja convertida em uma distância de mapa corrigida para múltiplos crossings. Suponha que, em um cruzamento-teste, obtemos um valor de FR de 27,5% (0,275). Substituindo-o na função, podemos resolver m :

$$0,275 = \frac{1}{2} (1 - e^{-m})$$

logo,

$$e^{-m} = 1 - (2 \times 0,275) = 0,45$$

Usando uma calculadora, podemos deduzir que $m = 0,8$. Isto é, em média, existem 0,8 crossings por meiose nessa região cromossômica.

A etapa final é converter essa medida de frequência de crossing para dar uma distância de mapa corrigida. O que temos que fazer para corrigir em unidades de mapa corrigida é multiplicar a frequência média de crossing calculada por 50, pois, em média, um crossing produz uma frequência de recombinantes de 50%. Assim, no exemplo numérico precedente, o valor m de 0,8 pode ser convertido em uma fração corrigida de recombinantes de $0,8 \times 50 = 40$ unidades corri-

gidas de mapa. Vemos que, de fato, esse valor é substancialmente maior que as 27,5 unidades de mapa que deduziríamos da FR observada.

Note que a função de mapa explica por que o valor FR máximo para genes ligados é de 50%. À medida que m se torna muito grande, e^{-m} tende a zero e a FR tende a 1/2 ou 50%.

A Fórmula de Perkins

Para fungos e outros organismos produtores de tétrade, existe um outro modo de compensar os múltiplos crossings, especificamente crossings duplos (o tipo esperado mais comum). Na análise de tétrade de “díbridos”, geralmente apenas três tipos de tétrade são possíveis, quando classificadas com base na presença de genótipos parentais e recombinantes nos produtos. De um cruzamento $AB \times ab$, eles são

Ditipo parental (DP)	Tetratipo (T)	Ditipo não-parental (DNP)
$A \cdot B$	$A \cdot B$	$A \cdot b$
$A \cdot B$	$A \cdot b$	$A \cdot b$
$a \cdot b$	$a \cdot B$	$a \cdot B$
$a \cdot b$	$a \cdot b$	$a \cdot B$

Os genótipos recombinantes são mostrados em vermelho. Se os genes estão ligados, um simples enfoque para mapear sua distância pode ser usado na seguinte fórmula:

$$\text{distância de mapa} = FR = 100 (DNP + \frac{1}{2} T)$$

porque essa fórmula dá a porcentagem de todos os recombinantes. Entretanto, nos anos de 1960, David Perkins desenvolveu uma fórmula que compensa os efeitos dos crossings duplos. A fórmula de Perkins fornece uma estimativa mais precisa da distância de mapa:

$$\text{distância de mapa corrigida} = 50 (T + 6 DNP)$$

Não faremos a derivação dessa fórmula, mas diremos que ela é baseada nos totais de classes DP, T e DNP esperados de meioses com 0, 1, e 2 crossings (ela supõe que números maiores são cada vez mais raros). Vejamos um exemplo de seu uso. Supomos que, em nosso cruzamento hipotético de $AB \times ab$, as frequências observadas de classes de tétrade são 0,56 DP, 0,41 T e 0,03 DNP. Usando a fórmula de Perkins, encontramos a distância de mapa corrigida entre os loci a e b como sendo

$$50[0,41 + (6 \times 0,03)] = 50(0,59) = 29,5 \text{ u.m.}$$

Comparemos esse valor com o valor não-corrigido obtido diretamente da FR. Usando os mesmos dados, encontramos:

$$\begin{aligned} \text{distância de mapa não-corrigida} &= 100 (\frac{1}{2} T + DNP) \\ &= 100 (0,205 + 0,03) \\ &= 23,5 \text{ u.m.} \end{aligned}$$

Essa distância é de 6 u.m. menor que a estimada que obtivemos usando a fórmula de Perkins, pois não corrigimos para crossings duplos.

À parte, que valores de DP, DNP e T são esperados quando lidamos com genes *não-ligados*? Os tamanhos de DP e DNP serão iguais como resultado da distribuição independente. A classe T pode ser produzida apenas por um crossing entre um dos dois loci e seus respectivos centrômeros, e, portanto, o tamanho da classe T dependerá do tamanho total das duas regiões que ficam entre o locus e o centrômero. Entretanto, a fórmula $\frac{1}{2} T + DNP$ sempre produz 0,50, refletindo a distribuição independente.

A tendência inerente de crossings múltiplos levarem a uma subestimação da distância de mapa pode ser evitada com o uso das funções de mapa (em qualquer organismo) e pela fórmula de Perkins (nos organismos produtores de tétrades, tais como fungos).

► 4.8 Uso de Mapas Baseados em Recombinação em Conjunto com Mapas Físicos

Os mapas de recombinação têm sido o principal tópico deste capítulo. Eles mostram os loci de *genes para os quais alelos mutantes (e seus fenótipos mutantes) foram encontrados*. As posições desse loci em um mapa são determinadas com base na frequência de recombinantes na meiose. A frequência de recombinantes é considerada proporcional à distância que separa os dois loci no cromossomo. Assim, a frequência de recombinantes torna-se a unidade de mapeamento. Tal mapeamento baseado em recombinação de genes com fenótipos mutantes conhecidos vem sendo feito há quase um século. Vimos como os sítios de heterozigose molecular (não-associados a fenótipos mutantes) também podem ser incorporados em tais mapas de recombinação. Como qualquer sítio heterozigoto, esses marcadores moleculares são mapeados por recombinação e, então, usados para navegar por um gene de interesse biológico. Fazemos a suposição perfeitamente razoável de que um mapa de recombinação representa o arranjo de genes nos cromossomos, mas, como já foi dito, esses mapas são construções realmente hipotéticas. Em contraste, os mapas físicos são o mais próximo do mapa genômico real que a ciência pode ter.

O tópico de **mapas físicos** será examinado mais atentamente no Capítulo 13, mas podemos dar uma idéia deles aqui. Um mapa físico é simplesmente um mapa do DNA genômico real, uma sequência muito longa de nucleotídeos, mostrando onde estão os genes, o quanto eles são grandes, o que há entre eles, além de outros marcos de interesse. As unidades de distância em um mapa físico são números de bases de DNA; por conveniência, o quilobase é a unidade preferida. A sequência completa de uma molécula de DNA é obtida sequenciando grandes números de pequenos fragmentos genômicos e, então, reunindo-os em uma sequência inteira. A sequência é, então, escaneada por um computador, programado para procurar segmentos similares a genes reco-

nhecidos por sequências de bases características, incluindo sequências de sinal conhecidas para a iniciação e término da transcrição. Quando o programa de computação encontra um gene, ele compara sua sequência com o banco de dados público de outros genes sequenciados para os quais foram descobertas funções em outros organismos. Em muitos casos, existe um “evento”; em outras palavras, a sequência assemelha-se muito a de um gene de função conhecida em outra espécie. Em tais casos, as funções de dois genes podem ser similares. A similaridade de sequência (em geral, próxima de 100%) é explicada pela herança do gene de algum ancestral comum e conservação geral de sequências funcionais ao longo do tempo evolutivo. Outros genes descobertos pelo computador não apresentam similaridade de sequência com nenhum gene de função conhecida. Assim, eles podem ser considerados “genes à procura de uma função”. Na realidade, é claro, é o pesquisador, não o gene, que procura e que deve encontrar a função. O sequenciamento de membros individuais diferentes de uma população também pode produzir sítios de heterozigose molecular, os quais, como fazem nos mapas de recombinação, atuam como marcadores de orientação no mapa físico.

Como os mapas físicos hoje estão disponíveis para a maioria dos organismos-modelo genéticos, existe de fato alguma necessidade de mapas de recombinação? Eles podem ser considerados fora de moda? A resposta é que ambos os mapas são usados em conjunto com outros para “triangular” na determinação da função gênica, um princípio ilustrado antes pelos mapas de Londres. O enfoque geral é ilustrado na Figura 4.25, que mostra um mapa físico e um mapa de recombinação da mesma região de um genoma. Ambos os mapas contêm genes e marcadores moleculares. Na parte inferior da Figura 4.25, vemos um trecho de um mapa baseado em recombinação, com posições de genes para os quais foram encontrados e mapeados fenótipos mutantes. Nem todos os genes nesse segmento estão incluídos. Para alguns desses genes, uma função pode ter sido descoberta com base em estudos bioquímicos ou outros de linhagens mutantes; genes para proteínas A e B, por exemplo. O gene no meio é um “gene de interesse” que um pesquisador descobriu afetar o aspecto do desenvolvimento que está sendo estudado. Para determinar sua função, o mapa físico pode ser útil. Os genes no mapa físico que estão na região geral de interesse no mapa de recombinação tornam-se *genes candidatos*, e qualquer um deles pode ser o gene de interesse. São necessários outros estudos para estreitar a escolha de um. Se esse único caso é um gene cuja função é conhecida para outros organismos, então é sugerida uma função para o gene de interesse. Desse modo, o fenótipo mapeado no mapa de recombinação pode ser associado a uma função deduzida do mapa físico. Os marcadores moleculares em ambos os mapas (não mostrados na Figura 4.25) podem ser alinhados para ajudar a centrar o processo. Assim, vemos que ambos os mapas contêm elementos de função: o mapa físico mostra a possível ação do gene a nível celular, enquanto o mapa de recombinação contém informação relacionada ao efeito

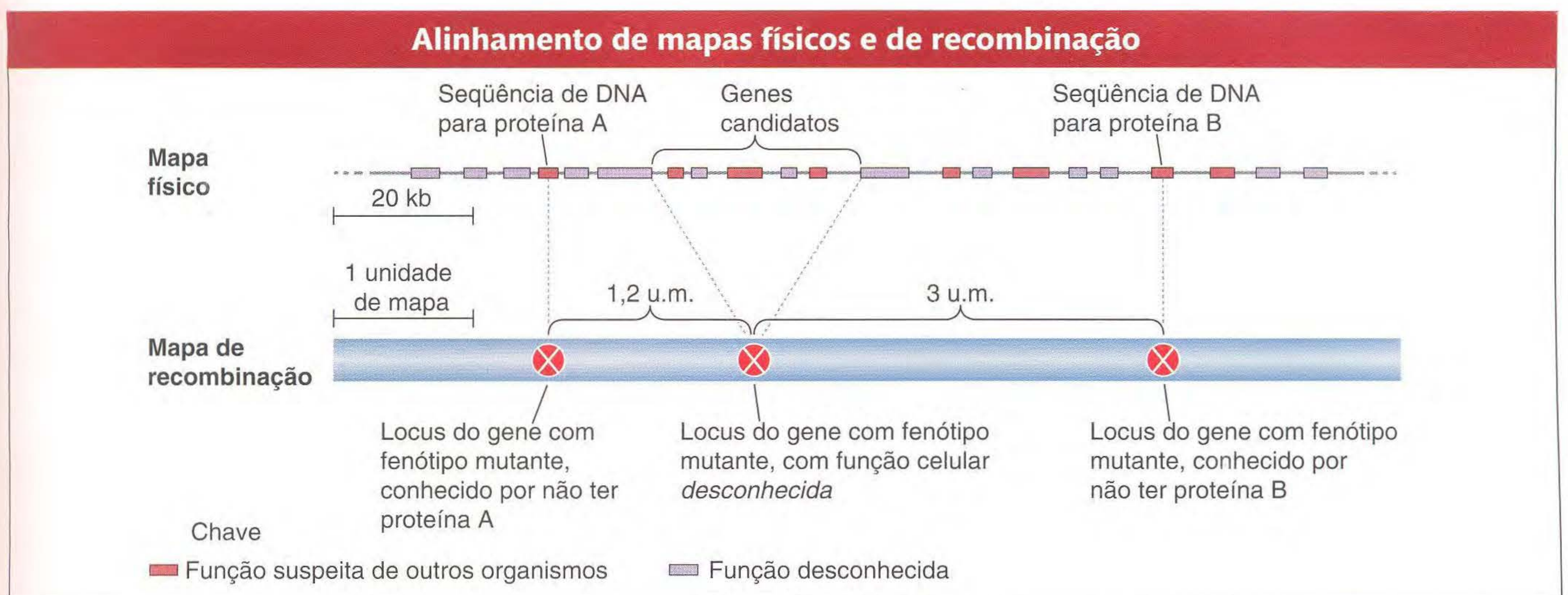


FIG. 4.25 Comparação das posições relativas nos mapas físico e de recombinação podem associar o fenótipo a uma função gênica desconhecida.

do gene a nível fenotípico. Em algum estágio, os dois têm que ser unidos para compreender a contribuição do gene para o desenvolvimento do organismo.

A união da recombinação com os mapas físicos pode atribuir uma função bioquímica a um gene identificado por seu fenótipo mutante.

Resumo

Em um cruzamento-teste diíbrido em *Drosophila*, Thomas Hunt Morgan encontrou um desvio da lei de Mendel da distribuição independente. Ele postulou que os genes estavam situados no mesmo par de cromossomos homólogos. Essa relação é chamada de ligação.

A ligação explica por que as combinações gênicas parentais ficam juntas, mas não como surgem as combinações recombinantes (não-parentais). Morgan postulou que, na meiose, pode haver uma troca física de partes do cromossomo por um processo agora chamado de crossing over. Um resultado da quebra física e reunião de partes do cromossomo, o crossing over ocorre no estágio de quatro cromátides da meiose. Assim, existem dois tipos de recombinação meiótica. A recombinação pela distribuição independente mendeliana resulta em uma frequência recombinante de 50%. O crossing over resulta em uma frequência de recombinantes geralmente menor que 50%.

À medida que Morgan estudava mais genes ligados, ele foi descobrindo muitos valores diferentes para a frequência de recombinantes e imaginou se esses valores correspondiam às distâncias reais entre genes em um cromossomo. Alfred Sturtevant, um estudante de Morgan, desenvolveu um método para determinar a distância entre genes em um mapa de ligação, com base na FR. O modo mais fácil de medir a FR é com um cruzamento-teste de um diíbrido ou triíbrido. Os valores de FR calculados como percentagens podem ser usados como unidades de mapa para construir um mapa cromossômico mostrando os loci dos genes analisados. Em fungos ascomicetos, os centrômeros também podem ser localizados no mapa medindo as frequências de segregação da segunda divisão.

Marcadores moleculares, tais como SNP e SSLP, também são usados para mapeamento cromossômico. As ilhas de SNP (haplótipos) são relativamente estáveis com o tempo, resultando em desequilíbrio de ligação. Haplótipos são usados para encontrar alelos mutantes porque os alelos SNP que ficam em desequilíbrio de ligação com um alelo mutante podem estar proximamente ligados.

Embora o teste básico para a ligação seja o desvio da distribuição independente, tal desvio pode não ser óbvio em um cruzamento-teste, e é necessário um teste estatístico. O teste do χ^2 , que diz com que frequência as observações se desviam das expectativas puramente por acaso, é particularmente útil em determinar se os loci estão ligados.

Os tamanhos das amostras no heredograma humano são muito pequenos para permitir o mapeamento, mas dados cumulativos, expressos como valores Lod (logaritmo das possibilidades), podem apoiar a hipótese de ligação. Os valores Lod são aditivos, e, assim, todas as medidas feitas em um intervalo de mapa específico podem ser somadas para produzir um valor Lod cumulativo.

Alguns crossings múltiplos podem resultar em cromátides não-recombinantes, levando a uma subestimação da distância de mapa com base na FR. A função de mapa, aplicável em qualquer organismo, corrige essa tendência. A fórmula de Perkins tem o mesmo uso na análise de tetrade de fungos.

Geralmente, em genética, o mapa baseado em recombinação de loci conferindo fenótipos mutantes é usado em conjunto com um mapa físico tal como a sequência completa do DNA, que mostra todas as sequências similares a genes. O conhecimento da posição do gene em ambos os mapas permite associar a função celular ao efeito do gene no fenótipo.

Termos-chave

centimorgan (cM)
coeficiente de coincidência (c.o.c.)
conformação cis
conformação trans
crossing over
cruzamento-teste de três pontos
desequilíbrio de ligação
distribuição de Poisson
fingerprint de DNA
frequência de recombinantes (FR)
função de mapa
genes ligados
haplótipo
hipótese nula

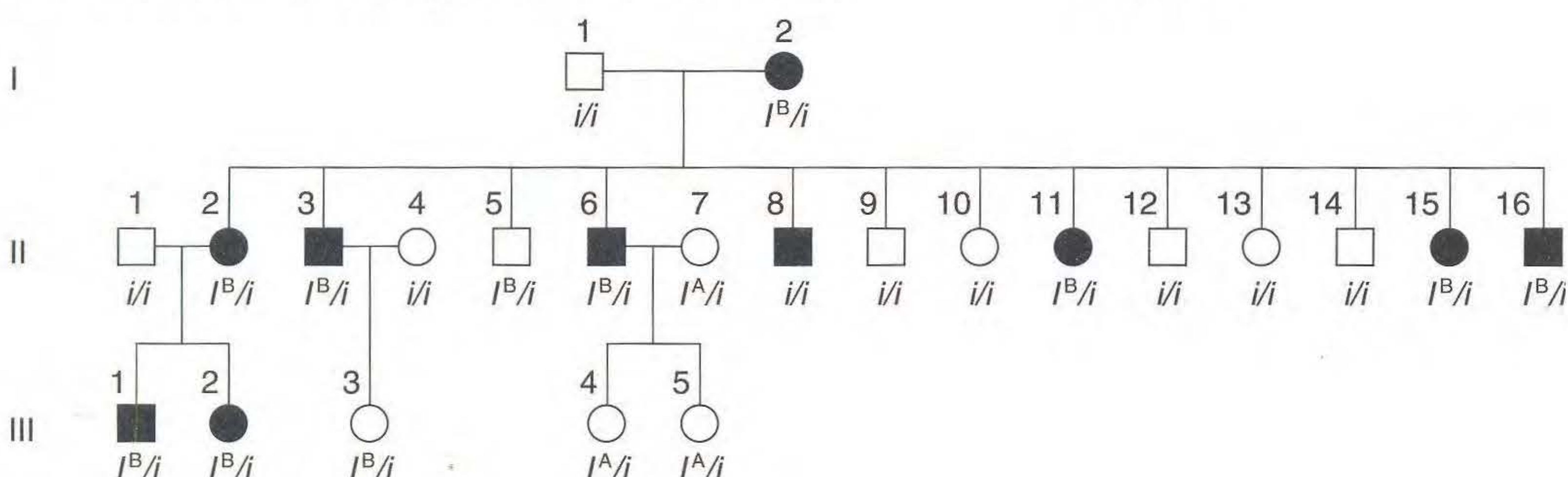
interferência
locus
mapa cromossômico
mapa de ligação
mapa de recombinação
mapa físico
marcador microssatélite
marcador minissatélite
marcador molecular
óctade
padrão de segregação de primeira divisão (padrão M_I)
padrão de segregação de segunda divisão (M_{II})

polimorfismo de comprimento de fragmento de restrição (RFLP)
polimorfismo de comprimento de sequência simples (SSLP)
polimorfismo de nucleotídeo único (SNP)
produto de crossing
repetições em tandem de número variável (VNTR)
unidade de mapa genético (u.m.)
valor Lod

Problemas Resolvidos

Problema resolvido 1. Um heredograma mostra uma pessoa afetada pela rara síndrome unha-rótula (unhas e joelhos malformados) e dá o genótipo do grupo sanguíneo ABO de cada pessoa. Ambos os loci envolvidos são autossômicos. Estude o seguinte heredograma:

- a. A síndrome unha-rótula é um fenótipo dominante ou recessivo? Cite os motivos para apoiar sua resposta.
b. Existe alguma evidência de ligação entre o gene unha-rótula e o gene do tipo sanguíneo ABO, julgando-se por este heredograma? Por que sim ou por que não?



SOLUÇÃO

- a. A síndrome unha-rótula mais provavelmente é dominante. Fomos informados de que ela é uma anomalia rara, e, assim, as pessoas não-afetadas que entram para a família improvavelmente levam um suposto alelo para síndrome unha-rótula. Seja N o alelo causador. Então, todas as pessoas com a síndrome serão heterozigotas N/n porque todas (provavelmente incluindo a avó) resultam de reproduções com pessoas normais n/n . Note que a síndrome aparece em todas as três gerações, outra indicação de herança dominante.
b. Há evidência de ligação. Note que a maioria das pessoas afetadas, que levam o alelo N , também levam o alelo I^B ; mais provavelmente, esses alelos estão ligados no mesmo cromossomo.

c.
$$\frac{n \quad i}{n \quad i} \times \frac{N \quad I^B}{n \quad i}$$

- c. Se há evidência de ligação, então desenhe os alelos nos homólogos relevantes dos avós. Se não há evidência de ligação, desenhe os alelos em dois pares de homólogos.
d. De acordo com seu modelo, que descendentes são recombinantes?
e. Qual a melhor estimativa de FR?
f. Se o homem III-1 se casa com uma mulher normal de tipo sanguíneo O, qual a probabilidade de que seu primeiro filho seja do grupo B com a síndrome unha-rótula?

(A avó deve levar ambos os alelos recessivos para produzir prole com o genótipo i/i e n/n .)

- d. Note que a reprodução dos avós é equivalente a um cruzamento-teste; logo os recombinantes na geração II são

$$II-5: n \ I^B/n \ i \text{ e } II-8: N \ i/n \ i$$

enquanto todos os outros são não-recombinantes, sendo ou $N \ I^B/n \ i$ ou $n \ i/n \ i$.

- e. Note que a reprodução dos avós e as duas primeiras reproduções na geração II são idênticas e são cruzamentos-teste. Três do total de 16 da prole são recombinantes (II-5, II-8 e III-3). A reprodução de II-6 com II-7 não é um cruzamento-teste, mas os cromossomos doados por II-6 podem ser deduzidos como sendo não-recombinantes. Assim, $FR = 3/18$, que é 17%.

f. (III-1 ♂) $\frac{N}{n} \frac{I^B}{i} \times \frac{n}{n} \frac{i}{i}$ (tipo O ♀ normal)

↓
Gametas

83,0%	{	NI^B	41,5%	← unha-rótula, tipo sanguíneo B
		ni	41,5%	
17,0%	{	Ni	8,5%	
		nI^B	8,5%	

As duas classes parentais são sempre iguais, e também as duas classes recombinantes. Logo, a probabilidade de que o primeiro filho tenha a síndrome unha-rótula e grupo sanguíneo B é de 41,5%.

Problema resolvido 2. O alelo b confere às moscas drosófilas um corpo preto, e b^+ dá uma cor marrom, o fenótipo tipo selvagem. O alelo wx de um gene separado dá asas graxas, e wx^+ dá não-graxas, o fenótipo tipo selvagem. O alelo cn de um terceiro gene dá olhos cinnabar, e cn^+ dá vermelho, o fenótipo tipo selvagem. Uma fêmea heterozigota para esses três genes é submetida a um cruzamento-teste, e uma prole de 1.000 é classificada do seguinte modo: 5 tipo selvagem; 6 pretas, graxas, cinnabar; 69 graxas, cinnabar; 67 pretas; 382 cinnabar; 379 pretas, graxas; 48 graxas; e 44 pretas, cinnabar. Note que o grupo da prole pode ser especificado listando apenas os fenótipos mutantes.

a. Explique esses números.

b. Desenhe os alelos em suas posições apropriadas nos cromossomos do triplo heterozigoto.

c. Se apropriado, de acordo com sua explicação, calcule a interferência.

SOLUÇÃO

a. Um alerta é ser metódico. Aqui, é uma boa idéia escrever os genótipos que podem ser deduzidos dos fenótipos. O cruzamento é um cruzamento-teste do tipo

$$b^+/b \cdot wx^+/wx \cdot cn^+/cn \times b/b \cdot wx/wx \cdot cn/cn$$

Note que existem pares distintos de classes de prole com relação à frequência. Já podemos calcular que as duas maiores classes representam cromossomos parentais, as duas classes com cerca de 68 representam crossings únicos em uma região, que as duas classes com cerca de 45 representam crossings únicos na outra região e que as duas classes com cerca de 5 representam crossings duplos. Podemos escrever a prole como classes derivadas dos gametas femininos, agrupados do seguinte modo:

$b^+ \cdot wx^+ \cdot cn$	382
$b \cdot wx \cdot cn^+$	379
$b^+ \cdot wx \cdot cn$	69
$b \cdot wx^+ \cdot cn^+$	67
$b^+ \cdot wx \cdot cn^+$	48
$b \cdot wx^+ \cdot cn$	44
$b \cdot wx \cdot cn$	6
$b^+ \cdot wx^+ \cdot cn^+$	5
	1.000

A listagem das classes desse modo confirma que os pares de classes são de fato genótipos recíprocos surgindo do zero, um ou dois crossings.

Primeiro, como não sabemos os genitores da fêmea tripla heterozigota, parece que não podemos aplicar a definição de recombinação na qual os genótipos gaméticos são comparados com dois genótipos parentais que formam uma mosca individual. Mas, refletindo, os únicos tipos parentais que fazem sentido com relação aos dados apresentados são $b^+/b^+ \cdot wx^+/wx^+ \cdot cn/cn$ e $b/b \cdot wx/wx \cdot cn^+/cn^+$, porque esses tipos são as classes gaméticas mais comuns.

Agora, podemos calcular as frequências de recombinantes. Para $b - wx$,

$$FR = \frac{69 + 67 + 48 + 44}{1.000} = 22,8\%$$

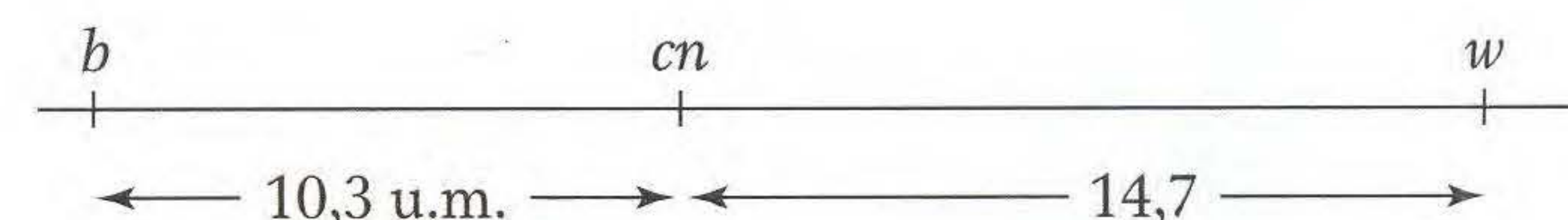
para $b - cn$,

$$FR = \frac{48 + 44 + 6 + 5}{1.000} = 10,3\%$$

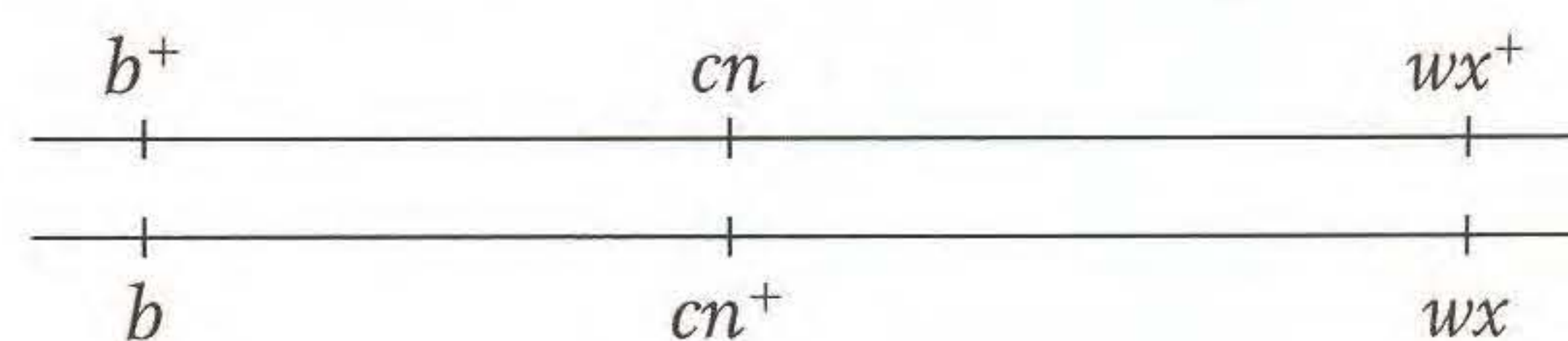
e para $wx - cn$,

$$FR = \frac{69 + 67 + 6 + 5}{1.000} = 14,7\%$$

O mapa é, portanto,



b. Os cromossomos parentais no heterozigoto triplo são



c. O número esperado de recombinantes duplos é $0,103 \times 0,147 \times 1.000 = 15,141$. O número observado é $6 + 5 = 11$; logo, a interferência pode ser calculada como

$$I = 1 - (11/15,141) = 1 - 0,726 = 0,274 = 27,4\%$$

Problema resolvido 3. É feito um cruzamento em linhagem haplóide de *Neurospora* com genótipo g^+ e outra linhagem haplóide com o genótipo g^- .

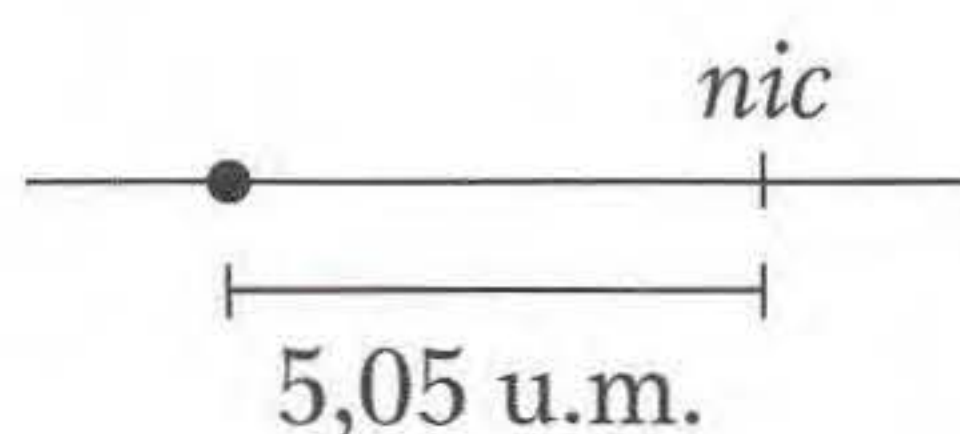
1	2	3	4	5	6	7
$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$	$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad^+$
$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$	$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad^+$
$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$	$nic \cdot ad$	$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad$
$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$	$nic \cdot ad$	$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad$
$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad^+$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$
$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad^+$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$	$nic^+ \cdot ad^+$	$nic^+ \cdot ad$
$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad$	$nic \cdot ad^+$	$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad^+$
$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad$	$nic \cdot ad^+$	$nic \cdot ad^+$	$nic \cdot ad$	$nic \cdot ad^+$
808	1	90	5	90	1	5

cruzamento, um total de 1.000 ascos lineares são isolados e categorizados como no quadro acima. Mapeie os loci *ad* e *nic* em relação aos centrômeros e um em relação ao outro.

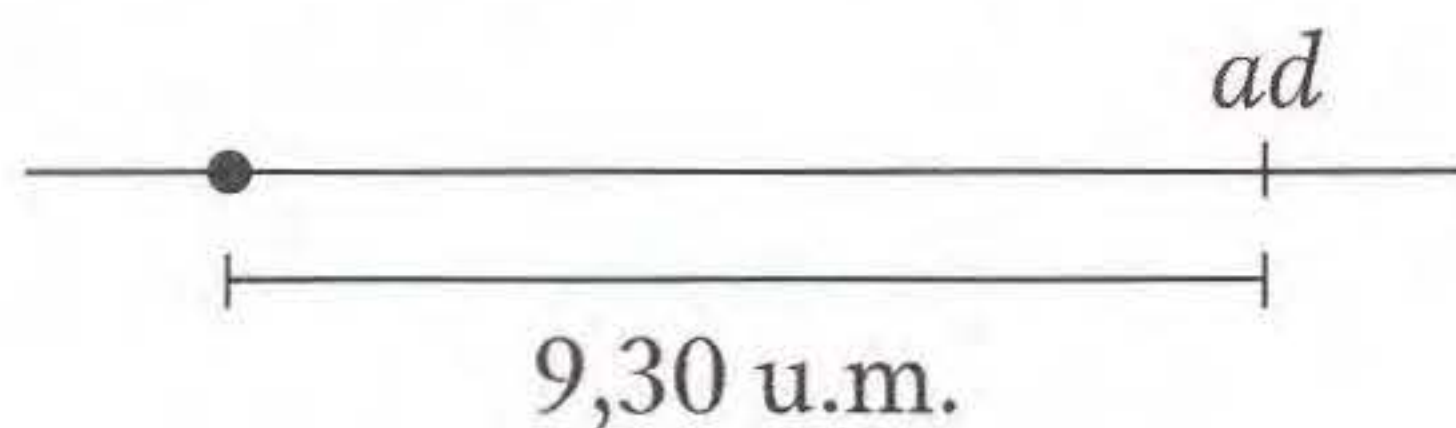
SOLUÇÃO

Que princípios podemos usar para resolver este problema? É uma boa idéia começar fazendo algo direto, que é calcular as distâncias dos dois locis ao centrômero. Não sabemos se os loci *ad* e *nic* estão ligados, mas não precisamos saber. As frequências dos padrões de M_{II} para cada locus dão a distância do locus para o centrômero. (Podemos nos preocupar sobre se é o mesmo centrômero depois.)

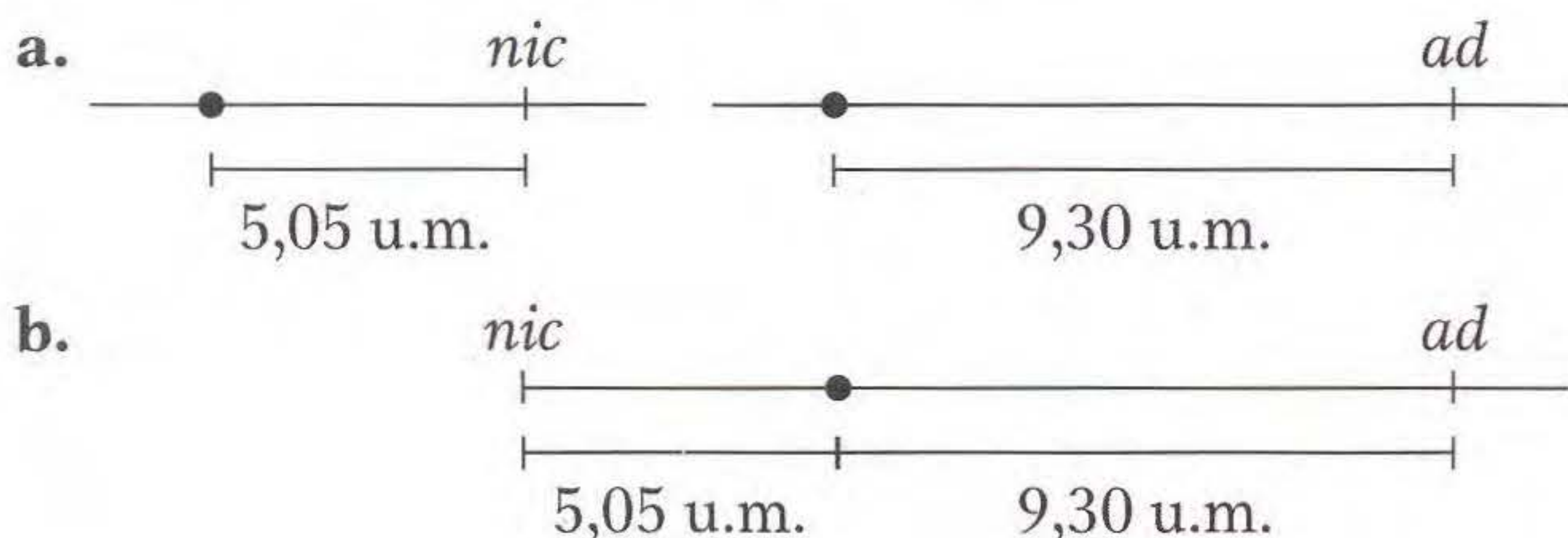
Lembre que um padrão M_{II} é qualquer padrão que não é dois blocos de quatro. Começemos com a distância entre o locus *nic* e o centrômero. Tudo que temos a fazer é somar os ascos tipos 4, 5, 6 e 7, pois todos eles são padrões de M_{II} para o locus *nic*. O total é $5 + 90 + 1 + 5 = 101$ de 1.000, ou 10,1%. Neste capítulo, já vimos que, para converter essa percentagem em unidades de mapa, temos que dividir por 2, o que dá 5,05 u.m.



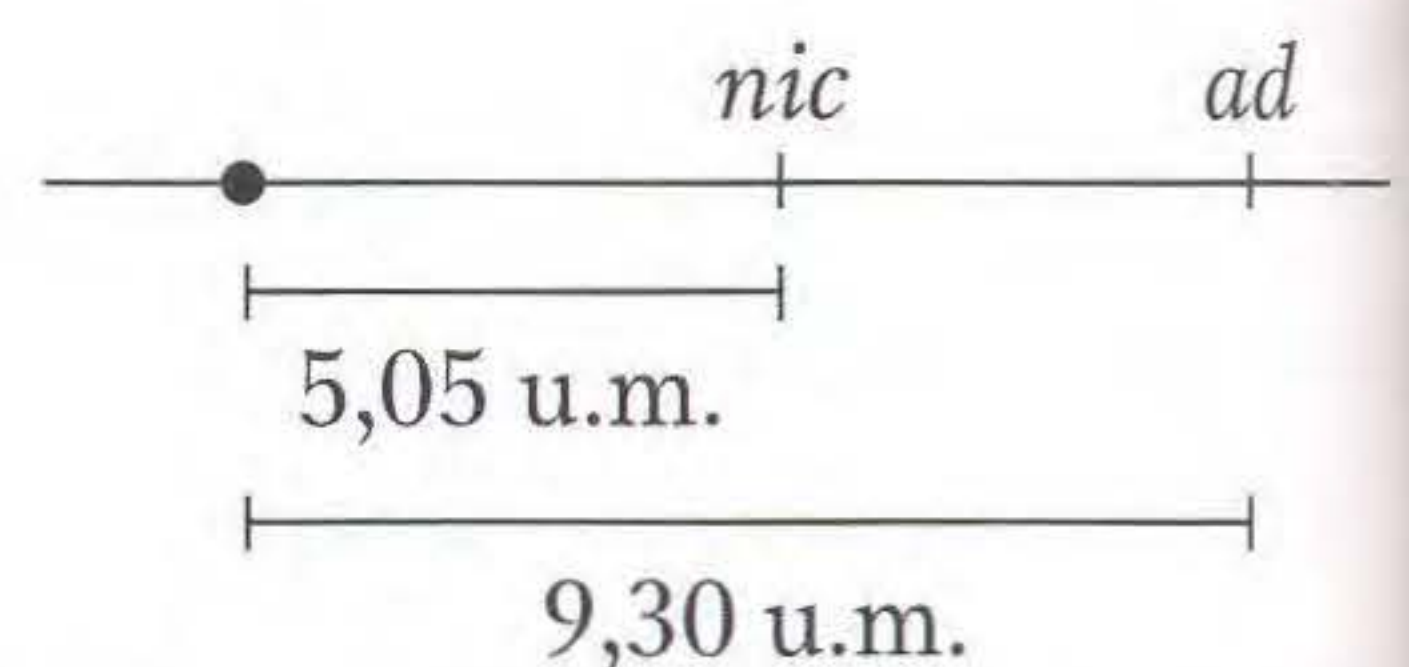
Fazemos o mesmo para o locus *ad*. Aqui, o total dos padrões M_{II} é dado pelos tipos 3, 5, 6 e 7 e é $90 + 90 + 1 + 5 = 186$ de 1.000, ou 18,6%, que é 9,3 u.m.



Agora, temos que juntar os dois e decidir entre as seguintes alternativas, todas as quais são compatíveis com as distâncias precedentes do locus ao centrômero:

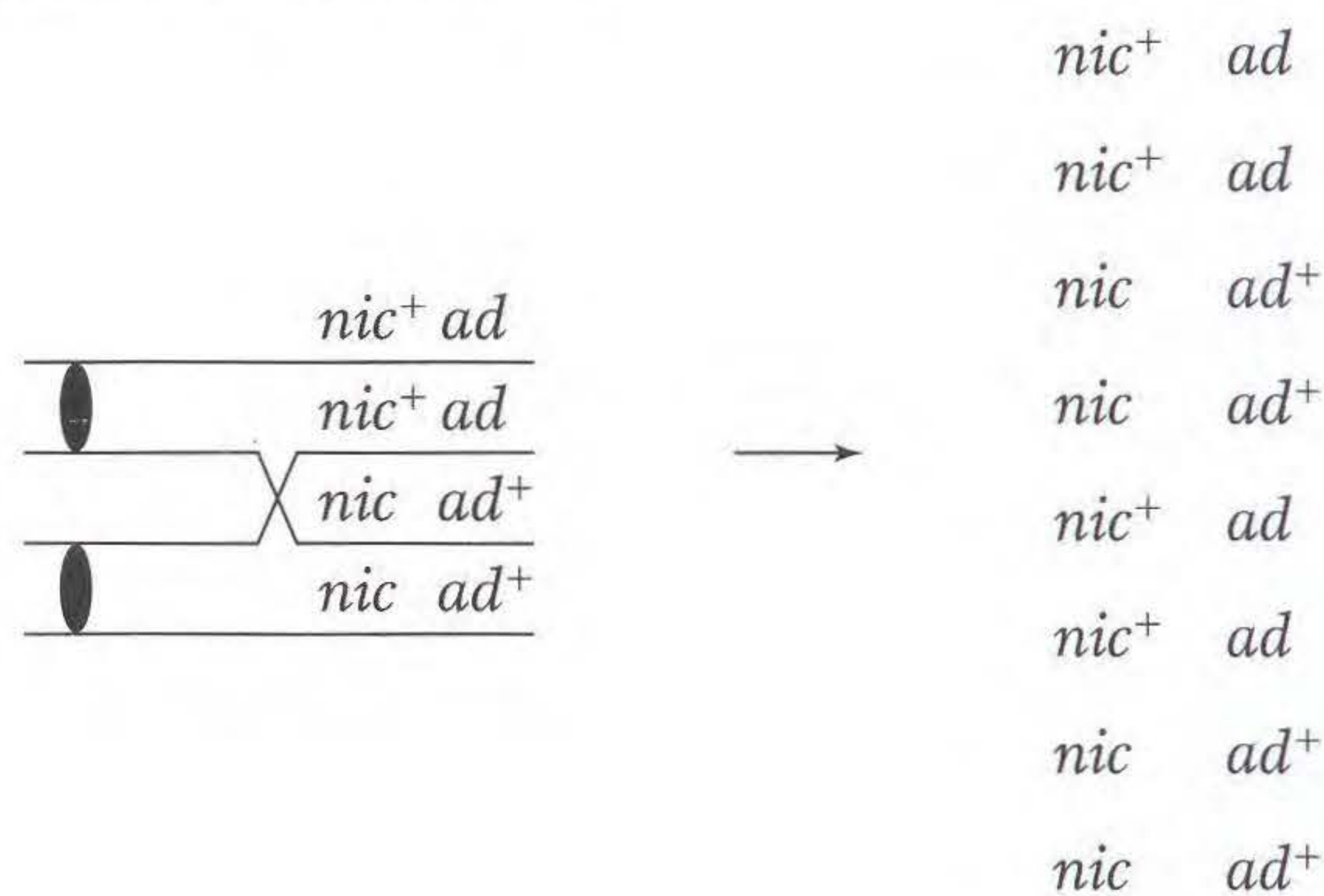


c.



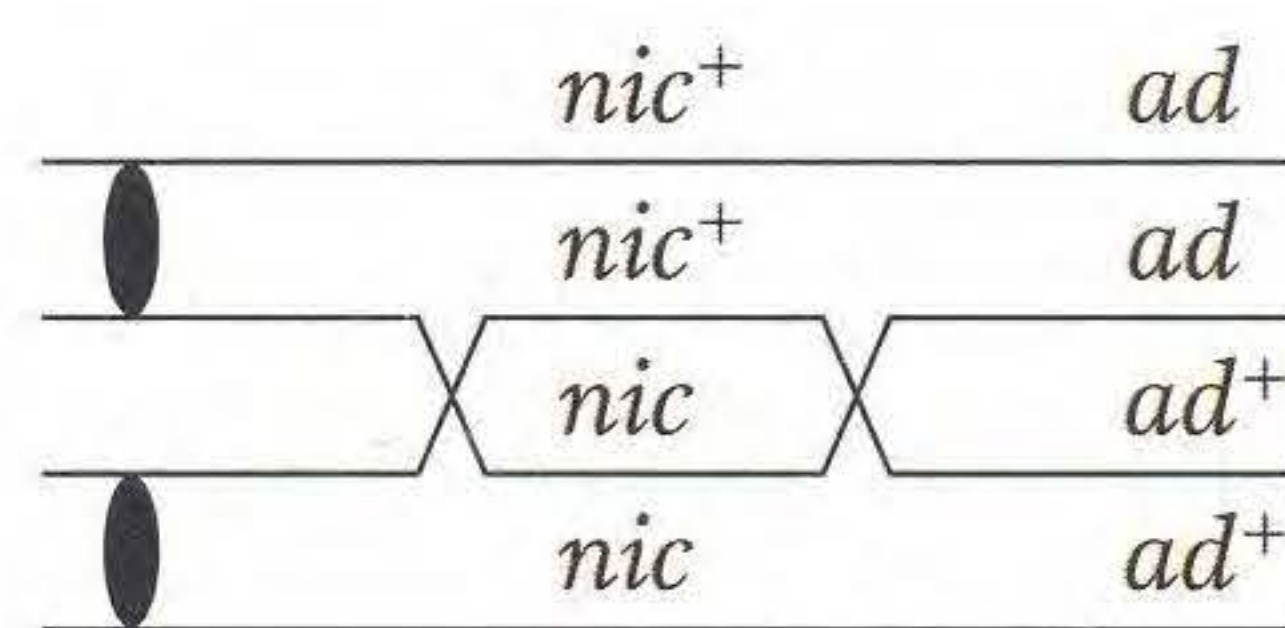
Aqui, uma combinação de bom senso e análise simples nos diz qual alternativa é a correta. Primeiro, uma inspeção dos ascos revela que o tipo mais comum é o marcado como 1, que contém 80% de todos os ascos. Esse tipo contém apenas os genótipos $nic^+ \cdot ad$ e $nic \cdot ad^+$, e eles são genótipos parentais. Assim, sabemos que a recombinação é bem baixa e que os loci estão certamente ligados. Isso elimina a alternativa a.

Agora, considere a alternativa c. Se essa alternativa fosse correta, um crossing entre o centrômero e o locus *nic* geraria não só um padrão M_{II} para esse locus, mas também um padrão M_{II} para o locus *ad*, porque ele está mais distante do centrômero do que *nic*. O padrão de asco produzido pela alternativa c deveria ser



Lembre que o locus *nic* apresenta padrões de M_{II} nos ascos tipo 4, 5, 6 e 7 (um total de 101 ascos); deles, o tipo 5 é aquele de que estamos falando e contém 90 ascos. Portanto, a alternativa c parece correta porque o asco tipo 5 compreende cerca de 90% dos ascos M_{II} para o locus *nic*. Essa relação não existiria se a alternativa b estivesse correta, pois os crossings em um dos lados do centrômero gerariam padrões M_{II} para os loci *nic* e *ad* independentemente.

A distância de mapa de *nic* para *ad* seria simplesmente $9,30 - 5,05 = 4,25$ u.m.? Perto, mas não tanto. O melhor modo de calcular as distâncias de mapa entre os loci é sem-



PDF FEITO PELA SSECP - UFRN - NATAL/RN

Classe	Fenótipos da prole	Gametas de F ₁	Números	Recombinante para		
				<i>p-b</i>	<i>p-v</i>	<i>v-b</i>
1	gre sen pla	+ · + · +	3.210			
2	pur res bro	<i>p · v · b</i>	3.222			
3	gre res pla	+ · <i>v · +</i>	1.024		R	R
4	pur sen bro	<i>p · + · b</i>	1.044		R	R
5	pur res pla	<i>p · v · +</i>	690	R		R
6	gre sen bro	+ · + · <i>b</i>	678	R		R
7	gre res bro	+ · <i>v · b</i>	72	R	R	
8	pur sen pla	<i>p · + · +</i>	60	R	R	
Total			10.000	1.500	2.200	3.436

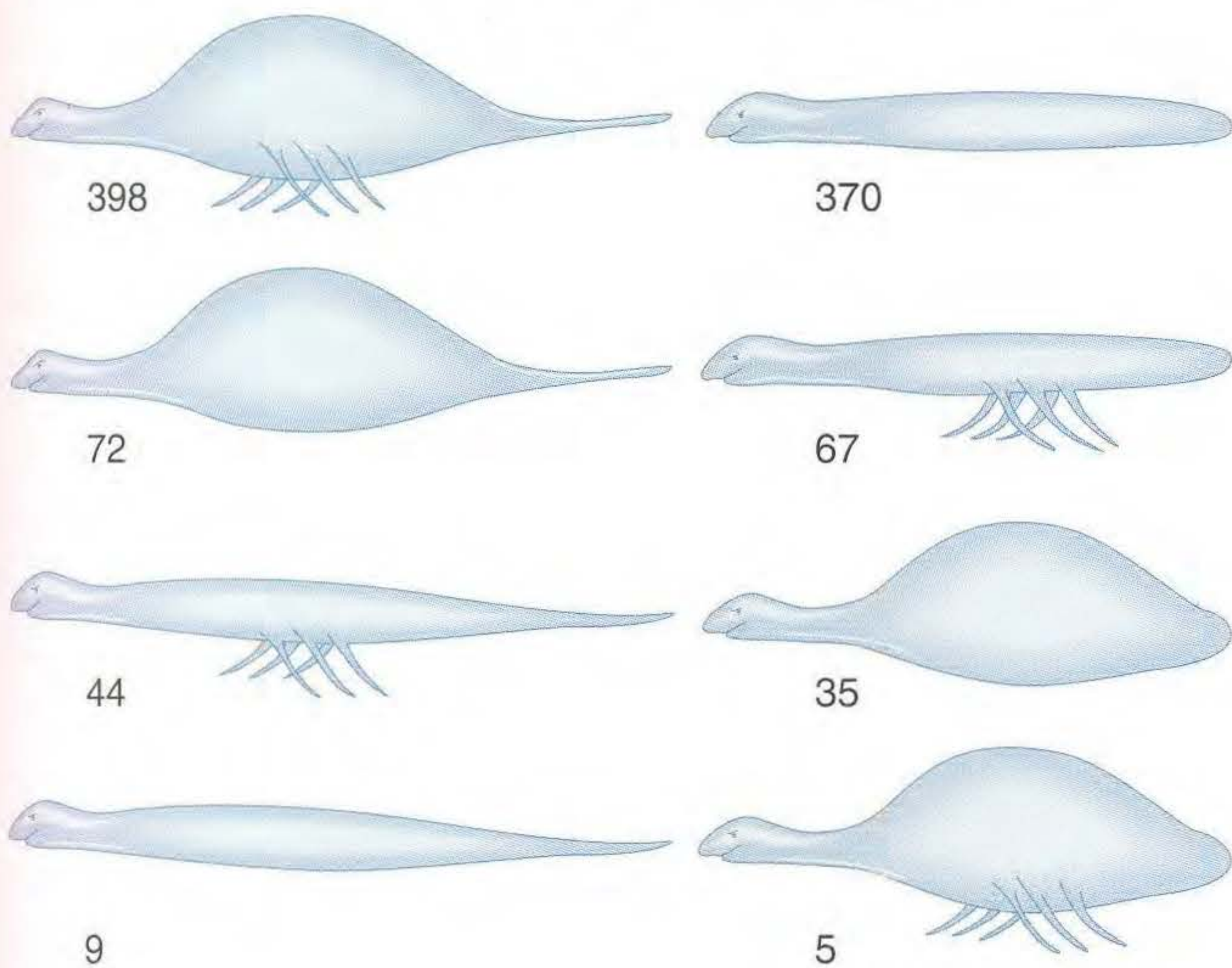


Solucionando o Problema 10

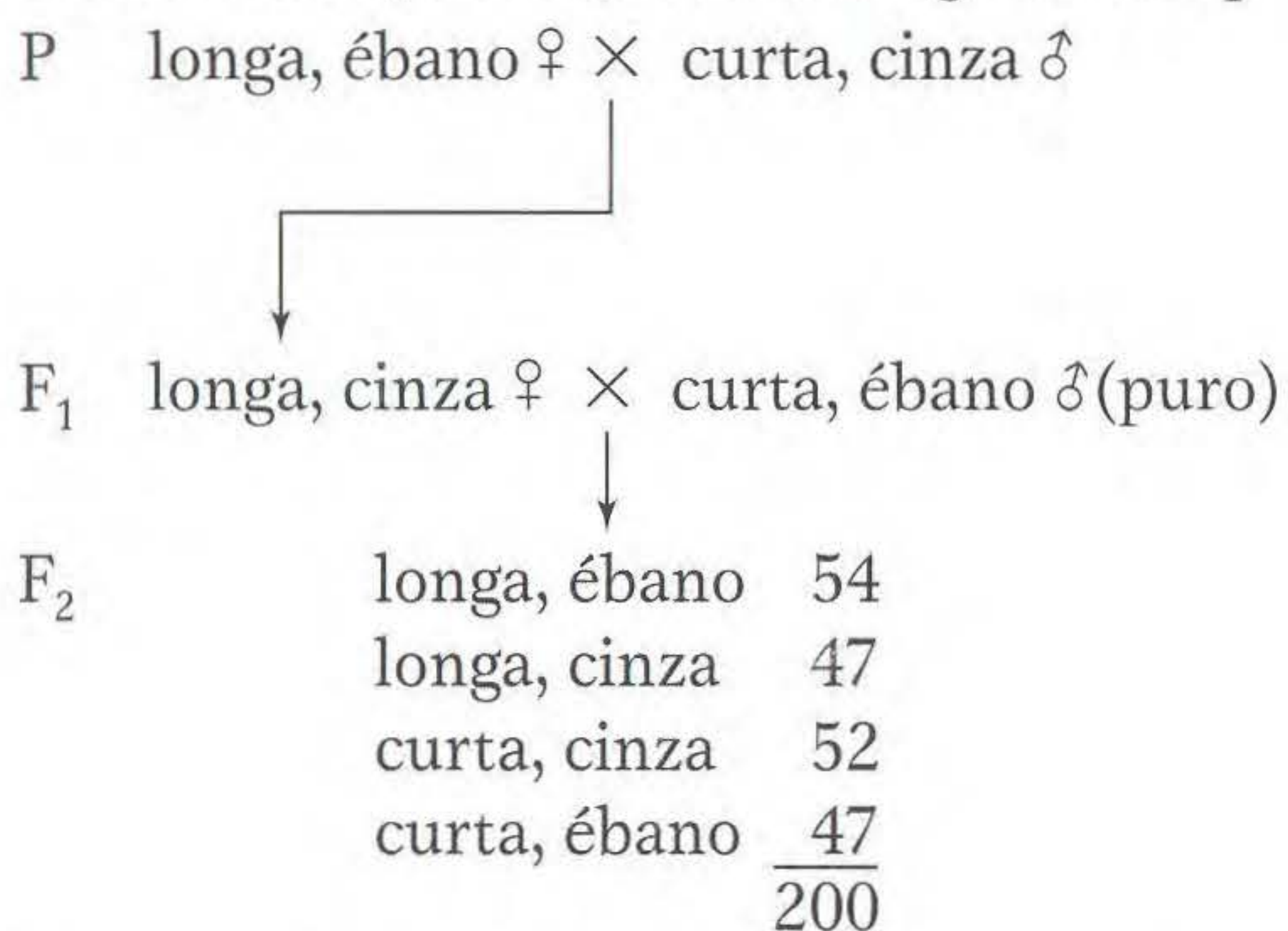
- Faça desenhos de P, F₁ e plantas testadoras de milho; use setas para mostrar exatamente como você faria esse experimento. Mostre onde as sementes são obtidas.
- Por que todos os + parecem iguais, mesmo para genes diferentes? Por que isso não causa confusão?
- Como pode um fenótipo ser púrpura e marrom, por exemplo, ao mesmo tempo?
- É significativo que os genes sejam escritos na ordem *p-v-b* no problema?
- O que é um testador e por que ele é usado na análise?
- O que a coluna marcada “Fenótipos da prole” representa? Na classe 1, por exemplo, diga exatamente o que significa “*gre sen pla*”.
- O que representa a linha marcada “Gametas”, e como ela é diferente da coluna marcada “gametas da F₁”? De que modo a comparação desses dois tipos de gametas é relevante para a recombinação?
- Que meiose é o principal foco de estudo? Marque-a em seu desenho.
- Por que os gametas do testador não são mostrados?
- Por que só existem oito classes fenotípicas? Estão faltando algumas classes?
- Que classes (e em que proporções) seriam esperadas se todos os genes estivessem em cromossomos separados?
- A que correspondem os quatro pares de tamanhos de classes (muito grande, duas intermediárias, muito pequena)?
- O que você pode dizer sobre a ordem dos genes simplesmente inspecionando as classes fenotípicas e suas frequências?
- Qual será a distribuição da classe fenotípica esperada se apenas dois genes estiverem ligados?
- O que a palavra “ponto” indica em um cruzamento-teste de três pontos? O uso dessa palavra significa ligação? Como seria um cruzamento-teste de quatro pontos?
- Qual a definição de *recombinante*, e como este é aplicado aqui?
- O que significam as colunas “Recombinante para”?
- Por que só existem três colunas “Recombinante para”?
- O que os R significam, e como eles são determinados?
- O que significa a coluna totais? Como os totais são usados?
- Qual o teste diagnóstico para ligação?
- O que é uma unidade de mapa? É o mesmo que um centimorgan?
- Em um cruzamento-teste como esse, por que a F₁ e o testador não são considerados como parentais no cálculo da recombinação? (Eles *são* genitores em um sentido.)
- Qual é a fórmula para a interferência? Como são calculadas as frequências esperadas na fórmula do coeficiente de coincidência?
- Por que a parte *c* do problema diz “se apropriado”?
- Quanto trabalho dá para obter uma prole de milho grande? Qual dos três genes dá mais trabalho para avaliar? Aproximadamente quanto de prole está representado em uma espiga de milho?
- Você tem uma linhagem de *Drosophila* que é homozigota para os alelos autossômicos recessivos *a*, *b* e *c*, ligados nessa ordem. Você cruza fêmeas dessa linhagem com machos homozigotos para as linhagens tipo selvagem correspondente. Você, então, cruza os machos heterozigotos de F₁ com suas irmãs heterozigotas. Você obtém os seguintes fenótipos F₂ (onde as letras indicam fenótipos recessivos e os sinais + indicam fenótipos tipo selvagem): 1.364 + + +, 365 *a b c*, 87 *a b* +, 84 + + *c*, 47 *a* + +, 44 + *b c*, 5 *a* + *c* e 4 + *b* +.
 - Qual a frequência de recombinantes entre *a* e *b*? Entre *b* e *c*? (Lembre-se de que não existem crossings nos machos de *Drosophila*.)
 - Qual é o coeficiente de coincidência?
- R.A. Emerson cruzou duas linhagens puras diferentes de milho e obteve uma F₁ fenotipicamente tipo selvagem e heterozigota para três alelos que determinam fenótipos recessivos: *an* determina antera; *br*, braquítico; e *f*, fino. Ele fez cruzamento-teste da F₁ com um testador que era homozigoto recessivo para três genes e obteve estes fenótipos de prole: 355 anteras; 339 braquíticos, finos; 88 completamente tipo selvagem; 55 anteras, braquíticos, finos; 21 finos; 17 anteras, braquíticos; 2 braquíticos; 2 anteras, finos.
 - Quais eram os genótipos das linhagens parentais?
 - Desenhe um mapa de ligação para os três genes (inclua as distâncias de mapa).
 - Calcule o valor de interferência.
- O cromossomo 3 de milho leva três loci (*b* para intensificador da cor da planta, *v* para virescência e *lg* para sem lígula). Um cruzamento-teste de triplos recessivos

com plantas F_1 heterozigotas para os três genes produz uma prole tendo os seguintes genótipos: 305 $+ v lg$, 275 $b + +$, 128 $b + lg$, 112 $+ v +$, 74 $+ + lg$, 66 $b v +$, 22 $+ + +$ e 18 $b v lg$. Escreva a seqüência gênica no cromossomo, as distâncias de mapa entre os genes e o coeficiente de coincidência.

14. Os groodies são úteis (mas fictícios) organismos haplóides que são puras ferramentas genéticas. Um groody tipo selvagem tem um corpo gordo, uma cauda longa e flagelos. As linhagens mutantes têm corpos finos, não têm cauda ou não têm flagelos. Os groodies podem se reproduzir uns com os outros (embora eles sejam tão envergonhados que não sabemos como) e produzem recombinantes. Um groody tipo selvagem reproduz-se com um groody magro que não tem cauda nem flagelo. Os 1.000 groodizinhos produzidos são classificados como mostrado na ilustração. Atribua genótipos e mapeie os três genes. (O Problema 14 é de Burton S. Guttman.)



15. Em *Drosophila*, o alelo dp^+ determina asas longas e dp determina asas curtas ("dumpy"). Em um locus separado, e^+ determina corpo cinza e e determina corpo ébano. Ambos os loci são autossômicos. Foram feitos os seguintes cruzamentos, começando com genitores puros:



Use o teste do χ^2 para determinar se esses loci estão ligados. Ao fazer isso, indique (a) a hipótese, (b) calcule o χ^2 , (c) o valor p , (d) o que o valor p significa, (e) sua conclusão, (f) a constituição cromossômica deduzida dos genitores, F_1 , testador e prole.

16. A mãe de uma família com 10 filhos tem o tipo Rh^+ . Ela também tem uma condição muito rara (eliptocitose, fenótipo E) que faz com que as hemácias sejam ovais em vez de apresentarem forma redonda, mas isso não produz efeitos clínicos adversos. O pai é Rh^- (falta de antígeno Rh^+) e tem hemácias normais (fenótipo e). As crianças são 1 Rh^+ e, 4 Rh^+ E e 5 Rh^- e. Há informações disponíveis sobre os genitores da mãe, que são Rh^+ E e Rh^- e. Um dos 10 filhos (que é Rh^+ E) casou-se com uma pessoa Rh^+ e, e tiveram um filho Rh^+ E.

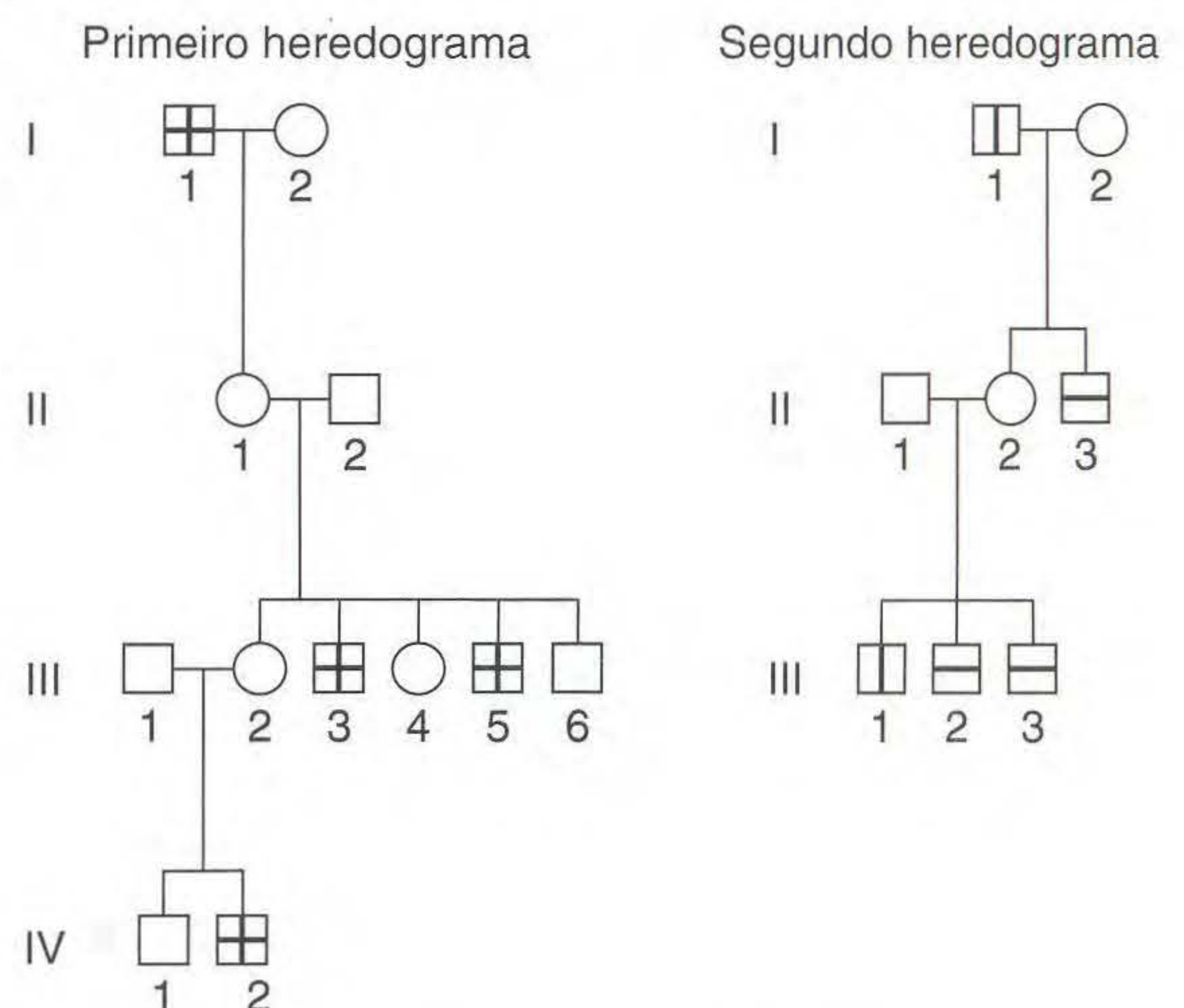
- a. Desenhe o heredograma de toda a família.
b. O heredograma está de acordo com a hipótese de que um alelo Rh^+ é dominante e Rh^- é recessivo?
c. Qual o mecanismo de transmissão da eliptocitose?
d. Os genes que controlam os fenótipos E e Rh poderiam estar no mesmo cromossomo? Caso sim, avalie a distância de mapa entre eles e comente seu resultado.

17. De vários cruzamentos do tipo geral $A/A \cdot B/B \times a/a \cdot b/b$, os indivíduos da F_1 do tipo $A/a \cdot B/b$ foram submetidos a um cruzamento-teste $a/a \cdot b/b$. Os resultados foram os seguintes:

Cruzamento-teste de F_1 do cruzamento	Prole de cruzamento-teste			
	$A/a \cdot B/b$	$a/a \cdot b/b$	$A/a \cdot b/b$	$a/a \cdot B/b$
1	310	315	287	288
2	36	38	23	23
3	360	380	230	230
4	74	72	50	44

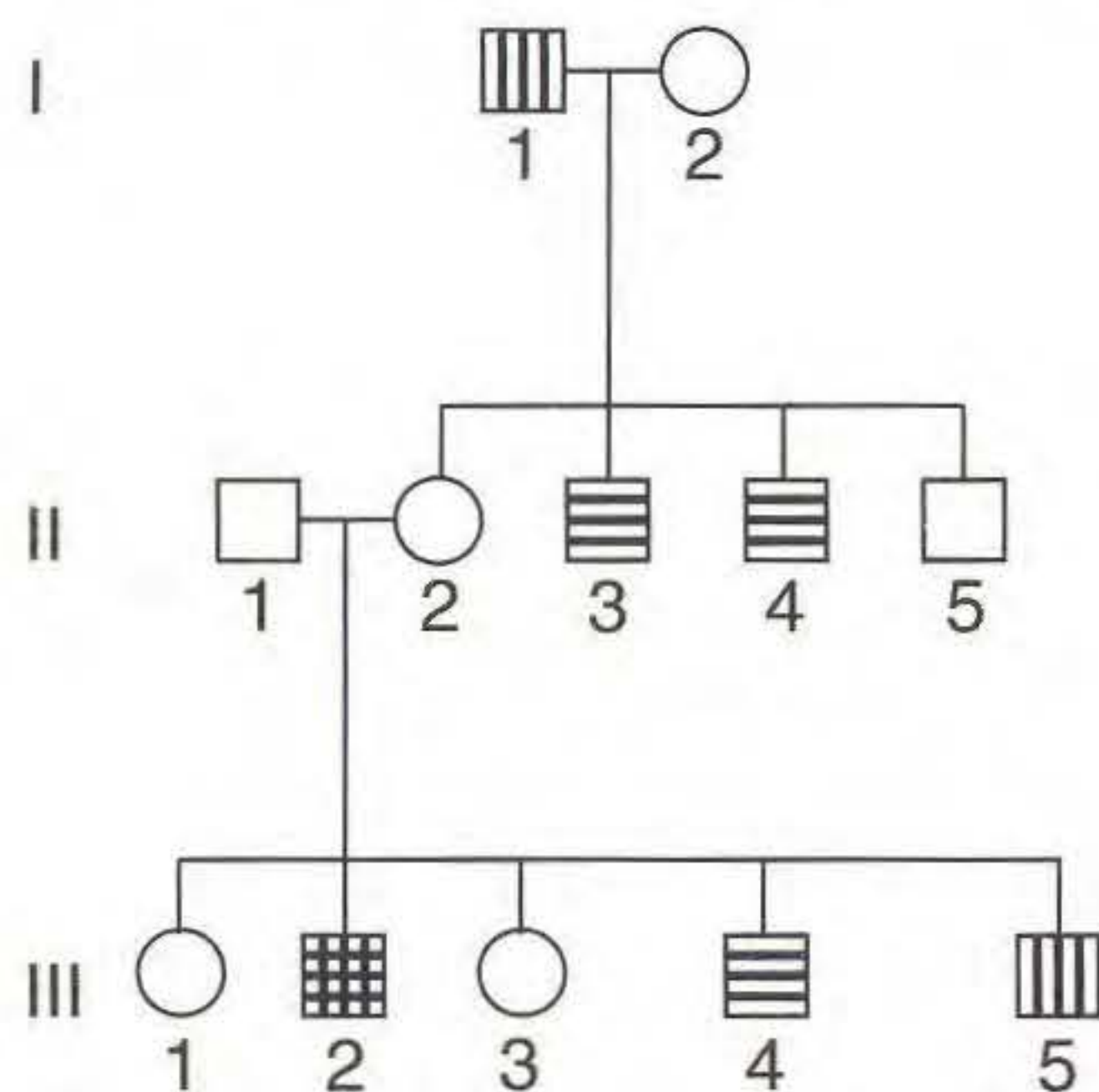
Para cada tipo de prole, use o teste do χ^2 para decidir se há evidência de ligação.

18. Nos dois heredogramas a seguir, uma barra vertical em um símbolo indica deficiência de esteróide sulfatase, e uma barra horizontal indica deficiência de ornitina transcarbamilase.



- a. Há alguma evidência nestes heredogramas de que os genes determinantes das deficiências são ligados?
b. Se os genes estão ligados, há alguma evidência no heredograma de crossing over entre eles?

- c. Desenhe os genótipos desses indivíduos tanto quanto possível.
19. No heredograma acompanhante, as linhas verticais indicam daltonismo protan, e as linhas horizontais indicam daltonismo deutan. Essas são condições separadas causando diferentes erros de percepção de cores; cada uma é determinada por um gene separado.



- a. O heredograma mostra alguma evidência de que os genes estão ligados?
- b. Se houver ligação, o heredograma mostra alguma evidência de crossing over?
- Explique suas respostas para as partes a e b com a ajuda do diagrama.
- c. Você pode calcular um valor para a recombinação entre esses genes? Essa recombinação é por distribuição independente ou por crossing over?
20. Em milho, um heterozigoto triplo foi obtido levando os alelos mutantes *s* (murcho), *w* (aleurona branca) e *y* (endosperma graxo), todos pareados com seus alelos tipo selvagem normais. Esse heterozigoto triplo foi submetido a cruzamento-teste, e a prole contém 116 murchos, brancos; 4 tipo selvagem; 2.538 murchos; 601 murchos, graxos; 626 brancos; 2.708 brancos, graxos; 2 murchos, brancos, graxos; e 113 graxos.
- a. Determine se algum desses três loci são ligados e, caso sim, mostre as distâncias de mapa.
- b. Mostre o arranjo de alelos nos cromossomos do heterozigoto triplo usado no cruzamento-teste.
- c. Calcule a interferência, se apropriado.
21. a. Foi feito um cruzamento de camundongos $A/a \cdot B/b \times a/a \cdot b/b$ e na prole havia

$$25\% A/a \cdot B/b, 25\% a/a \cdot b/b, \\ 25\% A/a \cdot b/b, 25\% a/a \cdot B/b$$

Explique essas proporções com a ajuda de diagramas simplificados de meiose.

- b. Foi feito um cruzamento de camundongos $C/c \cdot D/d \times c/c \cdot d/d$, e na prole havia

$$45\% C/c \cdot d/d, 45\% c/c \cdot D/d, \\ 5\% c/c \cdot d/d, 5\% C/c \cdot D/d$$

Explique essas proporções com a ajuda de diagramas simplificados de meiose.

22. Na pequena planta-modelo *Arabidopsis*, o alelo recessivo *hyg* confere resistência à semente para a droga higromicina

na, e *her*, um alelo recessivo de um gene diferente, confere resistência a herbicida. Uma planta que é homozigota *hyg/hyg \cdot her/her* foi cruzada com um tipo selvagem, e a F_1 foi autofecundada. As sementes resultantes da F_1 autofecundada foram colocadas em placas de Petri contendo higromicina e herbicida.

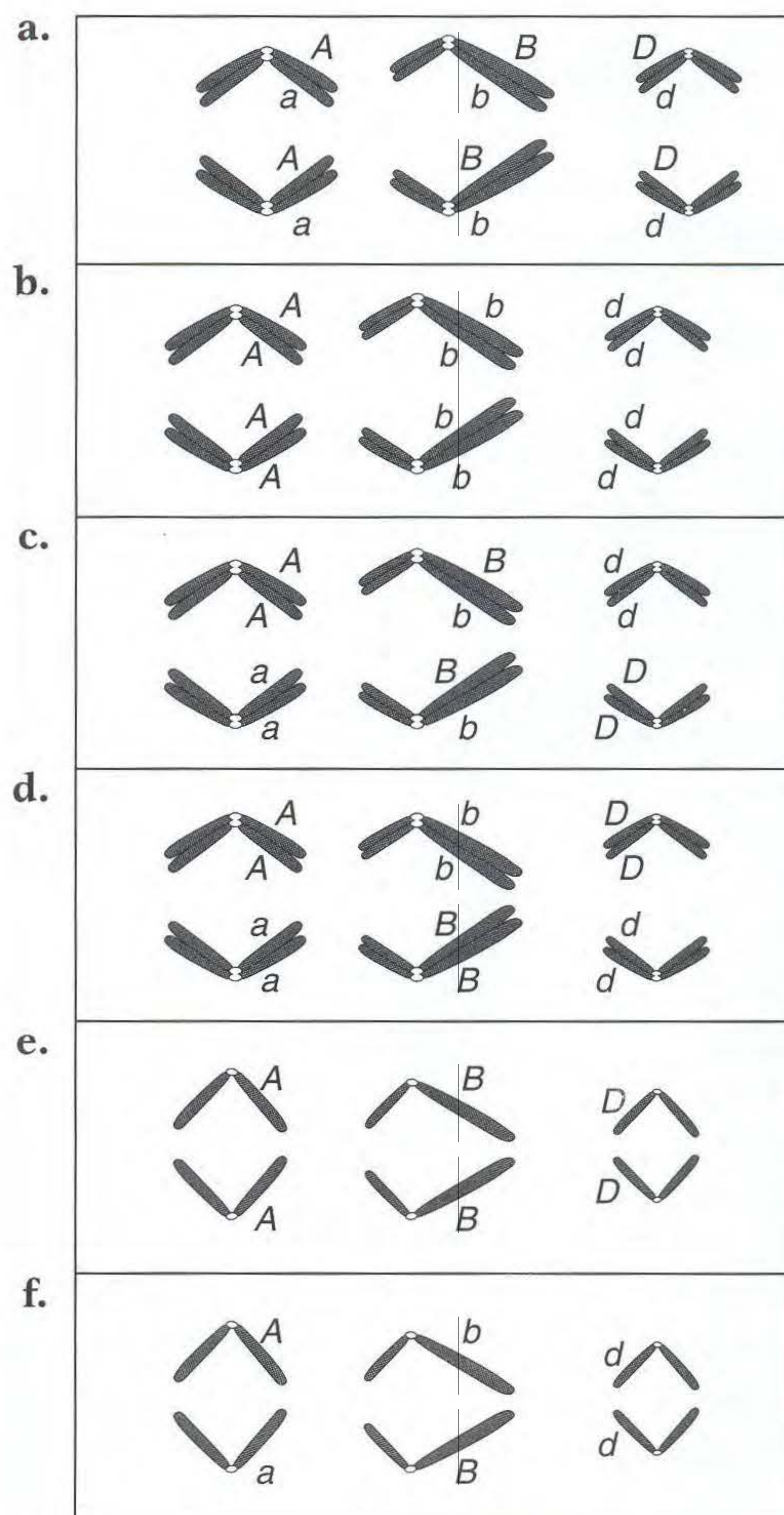
- a. Se os dois genes não estão ligados, qual a porcentagem de sementes que deverão crescer?

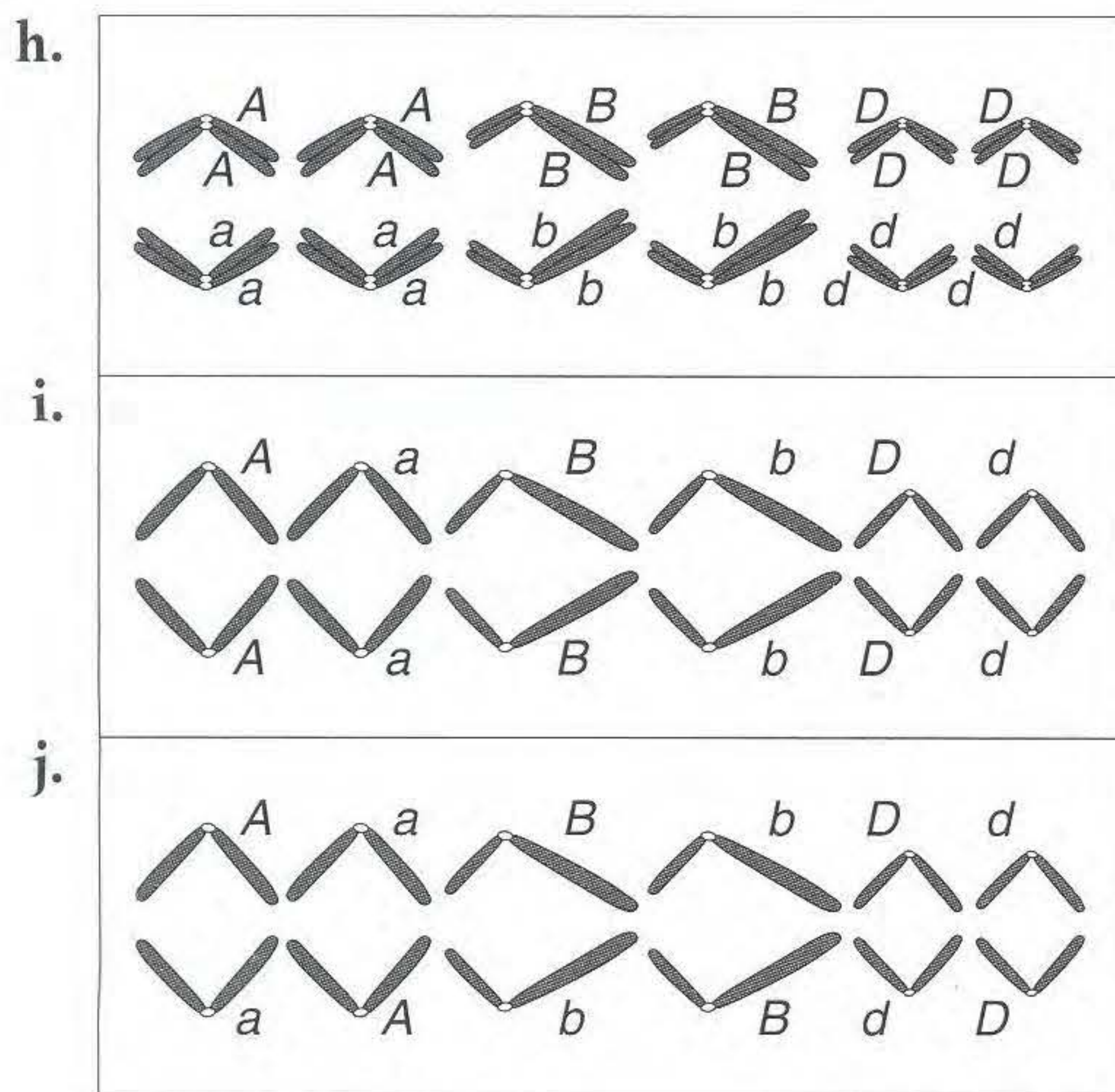
- b. De fato, 13% das sementes cresceram. Essa porcentagem apóia a hipótese de não-ligação? Explique. Caso não, calcule o número de unidades de mapa entre os loci.

- c. Em sua hipótese, se F_1 é submetida a cruzamento-teste, qual a proporção de sementes que crescerão em meio contendo higromicina e herbicida?

23. No heredograma da Figura 4.23, calcule o valor Lod para uma frequência de recombinantes de 34%.

24. Em um organismo diplóide com o genótipo $A/a \cdot B/b \cdot D/d$, os pares de alelos estão em diferentes pares de cromossomos. O diagrama a seguir mostra anáfases (estágios de "separação") em células individuais. Uma linha representa um cromossomo ou uma cromátide, e o ponto indica a posição do centrômero. Diga se cada desenho representa mitose, meiose I ou meiose II, ou se é impossível para esse genótipo em particular.





25. Foi feito um cruzamento de *Neurospora* $al-2^+ \times al-2$. Uma análise de tétrade linear revelou que a frequência de segregação da segunda divisão é de 8%.
- Desenhe dois exemplos de padrões de segregação de segunda divisão nesse cruzamento.
 - O que pode ser calculado usando o valor de 8%?

1	2	3	4	5	6	7
$A \cdot arg$	$A \cdot +$	$A \cdot arg$	$A \cdot arg$	$A \cdot arg$	$A \cdot +$	$A \cdot +$
$A \cdot arg$	$A \cdot +$	$A \cdot +$	$a \cdot arg$	$a \cdot +$	$a \cdot arg$	$a \cdot arg$
$a \cdot +$	$a \cdot arg$	$a \cdot arg$	$A \cdot +$	$A \cdot arg$	$A \cdot +$	$A \cdot arg$
$a \cdot +$	$a \cdot arg$	$a \cdot +$	$a \cdot +$	$a \cdot +$	$a \cdot arg$	$a \cdot +$
127	125	100	36	2	4	6



Solucionando o Problema 28

- Os fungos são geralmente haplóides ou diplóides?
- Quantos ascósporos estão no asco de *Neurospora*? Sua resposta corresponde ao número apresentado neste problema? Explique a discrepância.
- O que é tipo reprodutivo em fungos? Como você imagina que ele é determinado experimentalmente?
- Os símbolos A e a têm alguma coisa a ver com dominância e recessividade?
- O que significa o símbolo $arg-1$? Como você testaria esse genótipo?
- Como o símbolo $arg-1$ está relacionado ao símbolo $+$?
- O que significa a expressão *tipo selvagem*?
- O que significa a palavra *mutante*?
- A função biológica dos alelos mutantes tem algo a ver com a solução deste problema?
- O que significa a expressão *análise de óctade linear*?
- Em geral, o que mais pode ser aprendido da análise de tétrade linear que não pode ser aprendido da análise de tétrade desordenada?
- Como é feito um cruzamento em um fungo tal como *Neurospora*? Explique como isolar os ascos e ascósporos individuais. Como o termo tétrade se relaciona aos termos *asco* e *óctade*?

- Do cruzamento de fungos $arg-6 \cdot al-2 \times arg-6^+ \cdot al-2^+$, quais serão os genótipos dos esporos nas tétrades não-ordenadas que são (a) ditipos parentais? (b) tetratipos? (c) ditipos não-parentais?
- Para certa região cromossômica, o número médio de crossings na meiose é calculado como sendo de dois por meiose. Nessa região, qual a proporção de meioses previstas para ter (a) nenhum crossing? (b) um crossing? (c) dois crossings?
- Um cruzamento de *Neurospora* foi feito entre uma linhagem que levava o alelo A do tipo reprodutivo e o alelo mutante $arg-1$ e outra linhagem que levava o alelo tipo reprodutivo a e o alelo tipo selvagem para $arg-1$ ($+$). Foram isoladas 400 óctades lineares, e elas se enquadraram em sete classes do quadro a seguir. (Por simplificação, elas são mostradas como tétrades.)
 - Deduz o arranjo da linhagem do locus do tipo reprodutivo e o locus de $arg-1$. Inclua o centrômero ou centrômeros em qualquer mapa que você desenhe. Marque todos os intervalos em unidades de mapa.
 - Diagrama as divisões de meiose que levam à classe 6. Marque claramente.

- Onde ocorre a meiose no ciclo de vida de *Neurospora*? (Mostre em um diagrama do ciclo de vida.)
- O que o problema 28 tem a ver com meiose?
- Você pode escrever os genótipos das duas linhagens parentais?
- Por que são mostrados apenas quatro genótipos em cada classe?
- Por que existem apenas sete classes? Quantos modos você conhece para classificar tétrades em geral? Qual dessas classificações pode ser aplicada tanto a tétrades lineares quanto desordenadas? Você pode aplicar essas classificações a tétrades neste problema? (Classifique cada classe de tantos modos quanto possível.) Você pode pensar em mais possibilidades nesse cruzamento? Caso sim, por que elas não são mostradas?
- Você acha que existem várias ordens diferentes de esporos dentro de cada classe? Por que essas ordens diferentes de esporos não mudam a classe?
- Por que a seguinte classe não foi listada?

$a \cdot +$
 $a \cdot +$
 $A \cdot arg$
 $A \cdot arg$

20. O que significa a expressão *arranjo de ligação*?
21. O que é um *intervalo* genético?
22. Por que o problema diz “centrômero ou centrômeros” e não apenas “centrômero”? Qual é o método geral para mapeamento de centrômeros na análise de tétrades?
23. Qual é a frequência total de ascosporos $A \cdot +$? (Você calculou essa frequência usando uma fórmula ou por inspeção? Esse é um genótipo recombinante? Caso sim, ele é o único genótipo recombinante?)
24. As duas primeiras classes são as mais comuns e são aproximadamente iguais em frequência. O que essa informação lhe diz? Qual o seu conteúdo de genótipos parentais e recombinantes?
29. Um geneticista estuda 11 pares diferentes de loci de *Neurospora* fazendo cruzamentos do tipo $a \cdot b \times a^+ \cdot b^+$ e, então, analisa 100 ascas lineares de cada cruzamento. Para a conveniência de fazer um quadro, o geneticista organiza os dados como se todos os 11 pares de genes tivessem a mesma designação, a e b , mostrada a seguir. Para cada cruzamento, mapeie os loci em relação uns aos outros e aos centrômeros.

Número de ascas do tipo							
Cruzamento	$a \cdot b$	$a \cdot b^+$	$a \cdot b$	$a \cdot b$	$a \cdot b$	$a \cdot b^+$	$a \cdot b^+$
	$a \cdot b$	$a \cdot b^+$	$a \cdot b^+$	$a^+ \cdot b$	$a^+ \cdot b^+$	$a^+ \cdot b$	$a^+ \cdot b$
	$a^+ \cdot b^+$	$a^+ \cdot b$	$a^+ \cdot b^+$	$a^+ \cdot b^+$	$a^+ \cdot b^+$	$a^+ \cdot b$	$a^+ \cdot b^+$
	$a^+ \cdot b^+$	$a^+ \cdot b$	$a^+ \cdot b$	$a \cdot b^+$	$a \cdot b$	$a \cdot b^+$	$a \cdot b$
1	34	34	32	0	0	0	0
2	84	1	15	0	0	0	0
3	55	3	40	0	2	0	0
4	71	1	18	1	8	0	1
5	9	6	24	22	8	10	20
6	31	0	1	3	61	0	4
7	95	0	3	2	0	0	0
8	6	7	20	22	12	11	22
9	69	0	10	18	0	1	2
10	16	14	2	60	1	2	5
11	51	49	0	0	0	0	0

30. Três cruzamentos diferentes em *Neurospora* são analisados com base em tétrades desordenadas. Cada cruzamento combina um par diferente de genes ligados. Os resultados são mostrados no seguinte quadro:

Cruzamento	Genitores	Ditipos		
		parentais (%)	Tetra- tipos (%)	não- parentais (%)
1	$a \cdot b^+ \times a^+ \cdot b$	51	45	4
2	$c \cdot d^+ \times c^+ \cdot d$	64	34	2
3	$e \cdot f^+ \times e^+ \cdot f$	45	50	5

- Para cada cruzamento, calcule:
- a. frequência de recombinantes (FR).
- b. a distância de mapa não-corrigida, com base na FR.

- c. a distância de mapa corrigida, com base nas frequências de tétrades.
- d. a distância corrigida de mapa, com base na função de mapa.
31. No cromossomo 4 de *Neurospora*, o gene *leu3* está à esquerda do centrômero e sempre se segrega na primeira divisão, enquanto o gene *cys2* está à direita do centrômero e apresenta uma frequência de segregação de segunda divisão de 16%. Em um cruzamento entre uma linhagem *leu3* e uma linhagem *cis2*, calcule as seqüências previstas das seguintes sete classes de tétrades lineares, onde $l = leucina3$ e $c = cis2$. (Ignore os crossings duplos e outros múltiplos.)
- | | | | | | | |
|------------|-------------|--------------|-------------|------------|-------------|--------------|
| (i) $l\ c$ | (ii) $l\ +$ | (iii) $l\ c$ | (iv) $l\ c$ | (v) $l\ c$ | (vi) $l\ +$ | (vii) $l\ +$ |
| $l\ c$ | $l\ +$ | $l\ +$ | $+c$ | $++$ | $+c$ | $+c$ |
| $++$ | $+c$ | $++$ | $++$ | $++$ | $+c$ | $++$ |
| $++$ | $+c$ | $+c$ | $l\ +$ | $l\ c$ | $l\ +$ | $l\ c$ |
32. Um agricultor de arroz obteve um triplo heterozigoto levando os três alelos recessivos para flores albinas (al),

pragrana marrom (b) e folhas com penugem (fu), todos pareados com seus alelos normais tipo selvagem. Esse triplo heterozigoto foi submetido a um cruzamento-teste. Os fenótipos da prole eram

170	tipo selvagem
150	albinos, marrons, com penugem
5	marrons
3	albinos, com penugem
710	albinos
698	marrons, com penugem
42	com penugem
38	albinos, marrons

- a. Algum dos genes está ligado? Caso sim, desenhe um mapa marcado com as distâncias de mapa. (Não se preocupe com a correção para crossings múltiplos.)

valsador	cinza	inclinada	18
valsador	albino	inclinada	21
não-valsador	cinza	reta	19
não-valsador	albino	reta	22
valsador	cinza	reta	4
valsador	albino	reta	5
não-valsador	cinza	inclinada	5
não-valsador	albino	inclinada	6
Total			100

- a. Quais foram os genótipos dos dois camundongos parentais no cruzamento?
b. Desenhe os cromossomos nos genitores.
c. Se você deduziu ligação, cite o valor de unidade de mapa ou os valores e mostre como foram obtidos.

40. Considere o cruzamento de *Neurospora*

$$+; + \times f; p$$

É sabido que o locus $+/f$ está muito próximo do centrômero no cromossomo 7. De fato, tão próximo que nunca ocorrem segregações de segunda divisão. Também é conhecido que o locus $+/p$ está no cromossomo 5, a tal distância que geralmente há uma média de 12% de segregações de segunda divisão. Com essa informação, qual será a proporção de óctades que são

- a. ditipos parentais mostrando padrões M_I para ambos os loci?
b. ditipos não-parentais mostrando padrões M_I para ambos os loci?
c. tetratipos mostrando padrões M_I para $+/f$ e um padrão M_{II} para $+/p$?
d. tetratipos mostrando um padrão M_{II} para $+/f$ e um padrão M_I para $+/p$?

41. Em um fungo haplóide, os genes *al-2* e *arg-6* distam 30 unidades de mapa um do outro no cromossomo 1, e os genes *lis-5* e *met-1* distam 20 unidades de mapa no cromossomo 6. Em um cruzamento

$$al-2 +; + met-1 \times + arg-6; lis-5 +$$

que proporção da prole seria prototrófica $++; ++$?

42. Os alelos recessivos *k* (olhos em forma de rim em vez do tipo selvagem redondo), *c* (olhos de cor cardeal em vez do tipo selvagem vermelho) e *e* (corpo ébano em vez do cinza tipo selvagem) identificam três genes no cromossomo 3 de *Drosophila*. As fêmeas com olhos em forma de rim, cor cardeal, foram cruzadas com machos corpo ébano. A F_1 era tipo selvagem. Quando as fêmeas F_1 foram submetidas a cruzamento-teste com machos *kk cc ee*, foram obtidos os seguintes fenótipos de prole:

<i>k</i>	<i>c</i>	<i>e</i>	3
<i>k</i>	<i>c</i>	+	876
<i>k</i>	+	<i>e</i>	67
<i>k</i>	+	+	49
+	<i>c</i>	<i>e</i>	44
+	<i>c</i>	+	58
+	+	<i>e</i>	899
+	+	+	4
Total			2.000

- a. Determine a ordem dos genes e as distâncias de mapa entre eles.
b. Desenhe os cromossomos dos genitores e da F_1 .
c. Calcule a interferência e diga o que você pensa de seu significado.

PROBLEMAS DESAFIADORES

43. Dos genitores com os genótipos $A/A \cdot B/B$ e $a/a \cdot b/b$, foi produzido um diíbrido, a seguinte prole foi obtida na ordem indicada:

$$A/a \cdot B/b, a/a \cdot b/b, A/a \cdot B/b, A/a \cdot b/b, \\ a/a \cdot b/b, A/a \cdot B/b \text{ e } a/a \cdot B/b$$

Calcule o valor Lod desses resultados supondo valores de FR (a) 10%, (b) 20% e (c) 30%.

44. Use a função de mapa de Haldane para calcular a distância de mapa corrigida nos casos em que a FR medida = 5%, 10%, 20%, 30% e 40%. Faça um gráfico de FR contra a distância corrigida de mapa e use-o para responder à pergunta: Quando se deve usar a função de mapa?

45. Um indivíduo heterozigoto para quatro genes, $A/a \cdot B/b \cdot C/c \cdot D/d$, é submetido a um cruzamento-teste com $a/a \cdot b/b \cdot c/c \cdot d/d$, e uma prole de 1.000 foi classificada pela contribuição gamética dos genitores heterozigotos do seguinte modo:

$a \cdot B \cdot C \cdot D$	42
$A \cdot b \cdot c \cdot d$	43
$A \cdot B \cdot C \cdot d$	140
$a \cdot b \cdot c \cdot D$	145
$a \cdot B \cdot c \cdot D$	6
$A \cdot b \cdot C \cdot d$	9
$A \cdot B \cdot c \cdot d$	305
$a \cdot b \cdot C \cdot D$	310

- a. Que genes estão ligados?
b. Se duas linhagens puras foram cruzadas para produzir o indivíduo heterozigoto, que genótipos teriam?
c. Desenhe um mapa de ligação dos genes ligados, mostrando a ordem e as distâncias em unidades de mapa.
d. Calcule um valor de interferência, se apropriado.

46. Um alelo autossômico *N* em humanos causa anomalias nas unhas e nas rótulas, chamada síndrome unha-rótula. Considere os casamentos nos quais um genitor tem a síndrome unha-rótula e o tipo sanguíneo A, e o outro parceiro tem unhas e rótulas normais e sangue tipo O. Esses casamentos produzem alguns filhos que têm tanto síndrome unha-rótula quanto sangue tipo A. Suponha que filhos não-aparentados desse grupo fenotípico cresçam, se casem e tenham filhos. São observados quatro fenótipos nas seguintes percentagens nessa segunda geração:

síndrome unha-rótula, tipo sanguíneo A	66%
unhas e rótulas normais, tipo sanguíneo O	16%
unhas e rótulas normais, tipo sanguíneo A	9%
síndrome unha-rótula, tipo sanguíneo O	9%

Analise esses dados, explicando as frequências relativas dos quatro fenótipos. (Veja o Capítulo 6 para a base genética desses tipos sanguíneos.)

47. Suponha que três pares de alelos são encontrados em *Drosophila*: x^+ e x , y^+ e y e z^+ e z . Como mostrado pelos símbolos, cada alelo não-tipo selvagem é recessivo em relação a seu alelo tipo selvagem. Um cruzamento entre fêmeas heterozigotas e esses três loci e machos tipo selvagem produz uma prole tendo os seguintes genótipos: 1.010 fêmeas $x^+ \cdot y^+ \cdot z^+$, 430 machos $x \cdot y^+ \cdot z$, 441 machos $x^+ \cdot y \cdot z^+$, 39 machos $x \cdot y \cdot z$, 32 machos $x^+ \cdot y^+ \cdot z$, 30 machos $x^+ \cdot y^+ \cdot z^+$, 27 machos $x \cdot y \cdot z^+$, 1 macho $x^+ \cdot y \cdot z$ e 0 machos $x \cdot y^+ \cdot z^+$.

- a. Em que cromossomo de *Drosophila* estão esses genes?
b. Desenhe os cromossomos relevantes na genitora, mostrando o arranjo dos alelos.
c. Calcule as distâncias de mapa entre os genes e o coeficiente de coincidência.
48. Dos cinco conjuntos de dados citados no quadro seguinte, determine a ordem dos genes por exame, isto é, sem calcular os valores de recombinação. Os fenótipos recessivos são representados por letras minúsculas, e os dominantes, por sinais $+$.

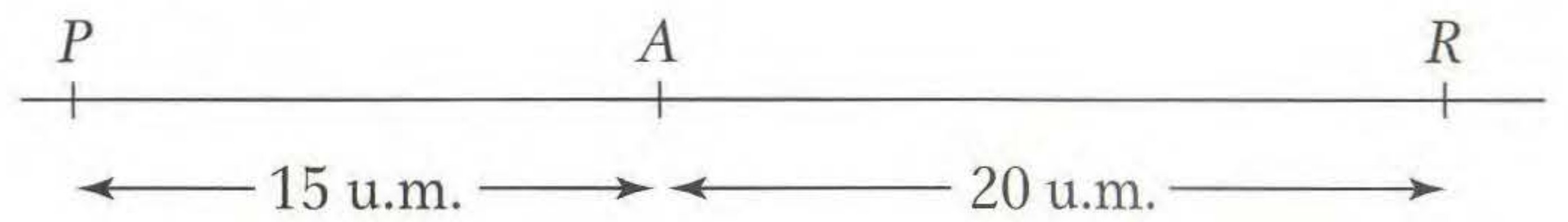
Fenótipos observados no cruzamento-teste de 3 pontos	Conjuntos de dados				
	1	2	3	4	5
$+$ $+$ $+$	317	1	30	40	305
$+$ $+$ c	58	4	6	232	0
$+$ b $+$	10	31	339	84	28
$+$ b c	2	77	137	201	107
a $+$ $+$	0	77	142	194	124
a $+$ c	21	31	291	77	30
a b $+$	72	4	3	235	1
a b c	203	1	34	46	265

49. Dos dados de fenótipos dados em seguida para cruzamentos-teste de três pontos para (1) a , b e c e (2) b , c e d , determine a sequência dos quatro genes a , b , c e d e as três distâncias de mapa entre eles. Os fenótipos recessivos são simbolizados por letras minúsculas e os genótipos dominantes por $+$.

1		2	
$+$ $+$ $+$	669	b c d	8
a b $+$	139	b $+$ $+$	441
a $+$ $+$	3	b $+$ d	90
$+$ $+$ c	121	$+$ c d	376
$+$ b c	2	$+$ $+$ $+$	14
a $+$ c	2.280	$+$ $+$ d	153
a b c	653	$+$ c $+$	65
$+$ b $+$	2.215	b c $+$	141

50. O pai de Mr. Spock, primeiro oficial da nave espacial *Enterprise*, veio do planeta Vulcano; a mãe de Spock veio da Terra. Um vulcano tem orelhas pontudas (determinado pelo alelo P), ausência de supra-renais (determinado por A) e um coração do lado direito (determinado por R). Todos esses alelos são dominantes em relação aos alelos da Terra. Os três loci são

autossômicos, e estão ligados como mostra o mapa de ligação:

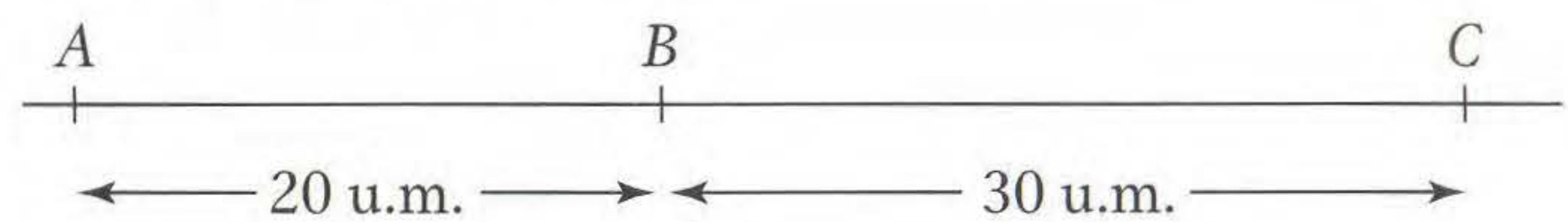


Se Mr. Spock se casar com uma mulher da Terra e não houver interferência genética, que proporção de seus filhos terão

- a. Fenótipos vulcanos para todas as três características?
b. Fenótipos terrestres para todas as três características?
c. Orelhas e coração vulcanos, mas supra-renais terrestres?
d. Orelhas vulcanas, mas coração e supra-renais terrestres?

(O Problema 50 é de D. Harrison, *Problems in Genetics*. Addison-Wesley, 1970.)

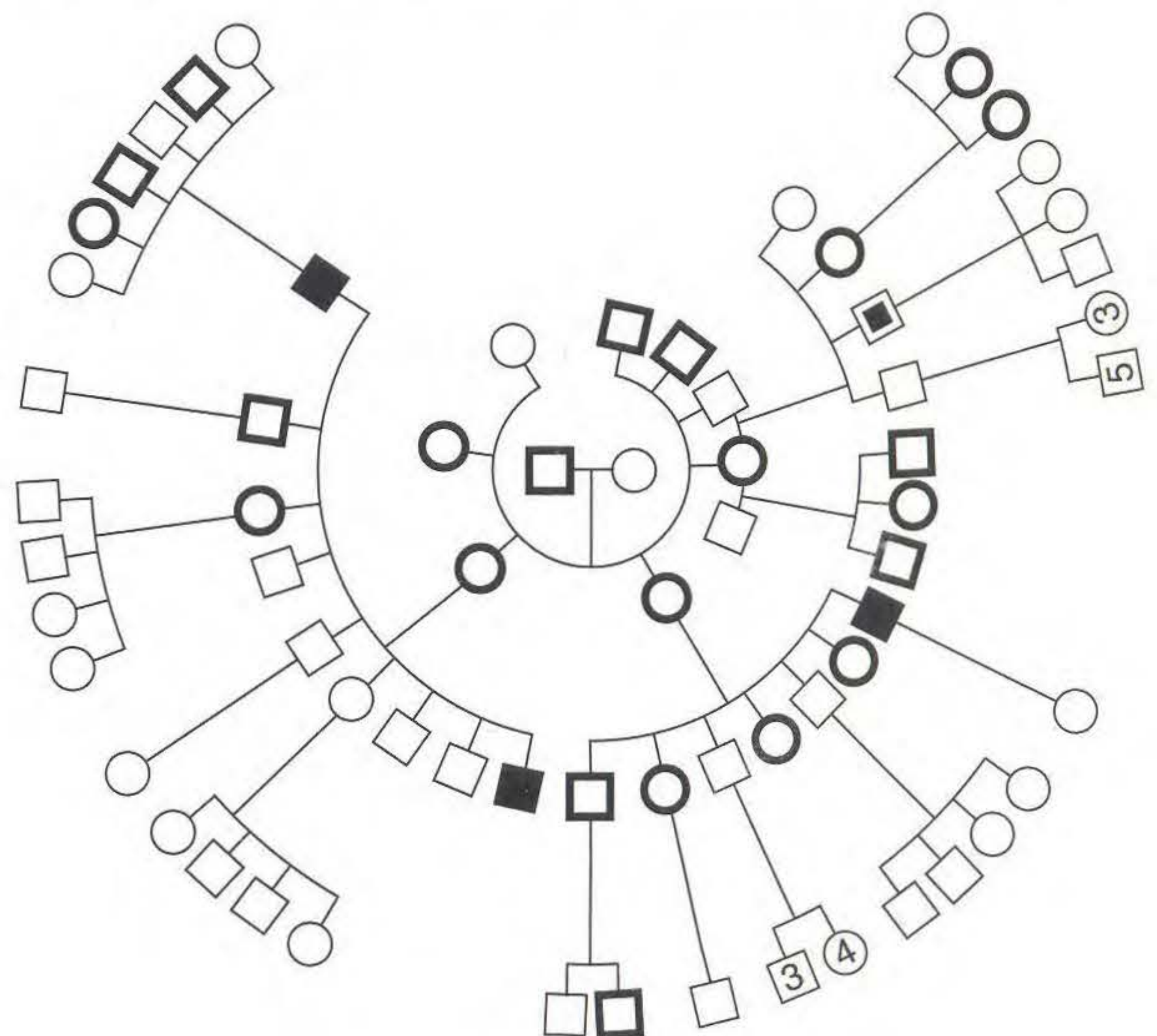
51. Em certa planta diplóide, os três loci A , B e C estão ligados do seguinte modo:



Uma planta está disponível para você (chame-a de planta parental). Ela tem a constituição $A b c/a B C$.

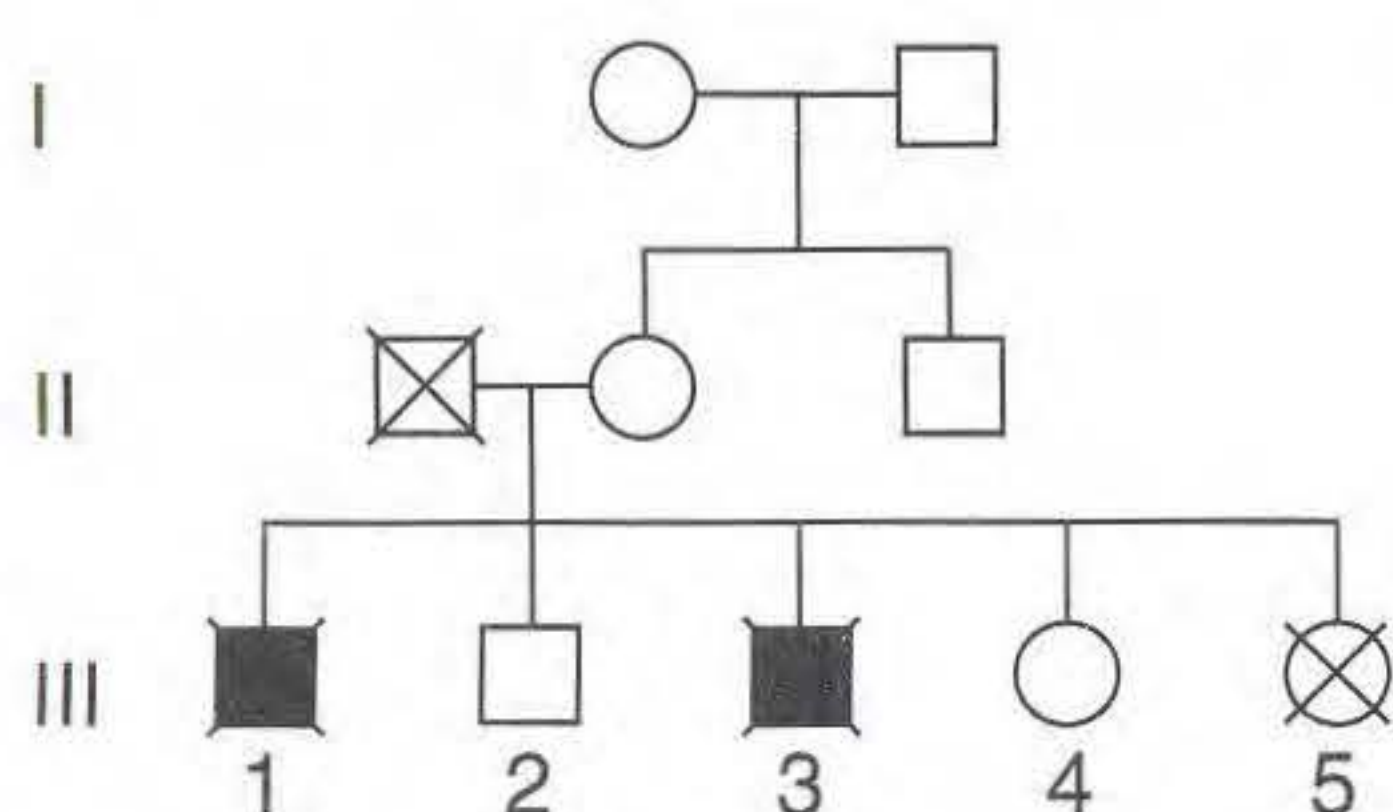
- a. Supondo não-interferência, se a planta for autofecundada, que proporção da prole terá o genótipo $a b c/a b c$?
b. Novamente, supondo não-interferência, se a planta genitora for cruzada com $a b c/a b c$, que classes genotípicas serão encontradas na prole? Quais serão suas frequências se houver uma prole de 1.000?
c. Repita a parte b, supondo agora 20% de interferência entre as regiões.

52. O heredograma seguinte mostra uma família com dois fenótipos anormais raros: escleróticas azuis (um defeito de ossos quebradiços), representado por um símbolo de borda escura, e hemofilia, representada por um centro escuro no símbolo. Os membros representados por símbolos totalmente pretos têm ambos os distúrbios. Os números em alguns símbolos são os números desses tipos.



- Que padrão de herança é mostrado por cada condição nesse heredograma?
- Cite os genótipos de tantos membros da família quanto possível.
- Há evidência de ligação?
- Há evidência de distribuição independente?
- Algum dos membros pode ser julgado como recombinante (isto é, formado de pelo menos um gameta recombinante)?

53. Os genes humanos para daltonismo e hemofilia estão ambos no cromossomo X, e eles mostram uma frequência de recombinantes de cerca de 10%. A ligação de um gene patológico com um relativamente não-prejudicial pode ser usada para prognóstico genético. Aqui é mostrada parte de um heredograma maior. Os símbolos escuros indicam que as pessoas têm hemofilia, e as cruzes indicam daltonismo. Que informação pode ser dada pelas mulheres III-4 e III-5 sobre a possibilidade de terem filhos com hemofilia?



(O problema 53 é adaptado de J.F. Crow, *Genetics Notes: An Introduction to Genetics*. Burgess, 1983.)

54. Um geneticista, mapeando os genes *A*, *B*, *C*, *D* e *E*, faz cruzamentos-teste de 3 pontos. O primeiro cruzamento de linhagens puras é

$$A/A \cdot B/B \cdot C/C \cdot D/D \cdot E/E \times a/a \cdot b/b \cdot C/C \cdot d/d \cdot E/E$$

Um geneticista cruza a F_1 com um testador recessivo e classifica a prole pela contribuição gamética da F_1 :

$A \cdot B \cdot C \cdot D \cdot E$	316
$a \cdot b \cdot C \cdot d \cdot E$	314
$A \cdot B \cdot C \cdot d \cdot E$	31
$a \cdot b \cdot C \cdot D \cdot E$	39
$A \cdot b \cdot C \cdot d \cdot E$	130
$a \cdot B \cdot C \cdot D \cdot E$	140
$A \cdot b \cdot C \cdot D \cdot E$	17
$a \cdot B \cdot C \cdot d \cdot E$	13
	1.000

O segundo cruzamento de linhagens puras é

$$A/A \cdot B/B \cdot C/C \cdot D/D \cdot E/E \times a/a \cdot B/B \cdot c/c \cdot D/D \cdot e/e$$

O geneticista cruza a F_1 desse cruzamento com uma irmã recessiva e obtém

$A \cdot B \cdot C \cdot D \cdot E$	243
$a \cdot B \cdot c \cdot D \cdot e$	237
$A \cdot B \cdot c \cdot D \cdot e$	62
$a \cdot B \cdot C \cdot D \cdot E$	58
$A \cdot B \cdot C \cdot D \cdot e$	155
$a \cdot B \cdot c \cdot D \cdot E$	165
$a \cdot B \cdot C \cdot D \cdot e$	46
$A \cdot B \cdot c \cdot D \cdot E$	34
	1.000

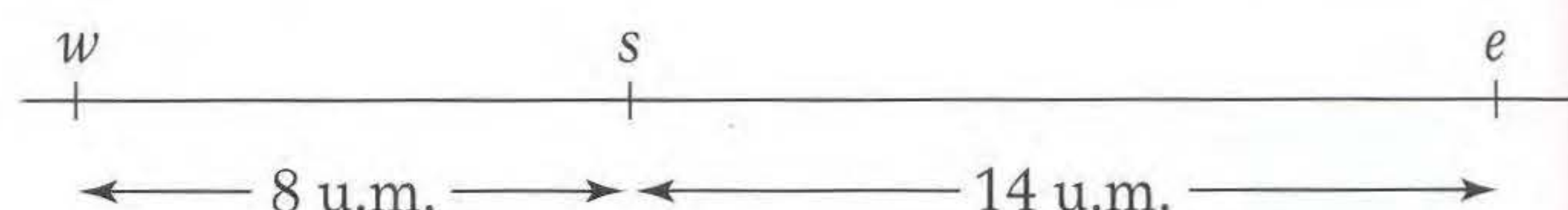
O geneticista também sabe que os genes *D* e *E* distribuem-se independentemente.

- Desenhe um mapa desses genes, mostrando as distâncias em unidades de mapa sempre que possível.
 - Há alguma evidência de interferência?
55. Na planta *Arabidopsis*, os loci para tamanho da vagem (*L*, longa; *l*, pequena) e pêlos nos frutos (*H*, piloso; *h*, liso) estão ligados distando 16 unidades de mapa no mesmo cromossomo. Foram feitos os seguintes cruzamentos:

$$\begin{aligned} \text{(i)} \quad & L H / L H \times l h / l h \rightarrow F_1 \\ \text{(ii)} \quad & L h / L h \times l H / l H \rightarrow F_1 \end{aligned}$$

Se as F_1 dos cruzamentos i e ii forem cruzadas,

- que proporção da prole é esperada como sendo $l h / l h$?
 - que proporção da prole é esperada sendo $L h / l h$?
56. Em milho (*Zea mays*), o mapa genético de parte do cromossomo 4 é o seguinte, onde *w*, *s* e *e* representam alelos mutantes recessivos que afetam a cor e a forma do pólen:



Se for feito o seguinte cruzamento

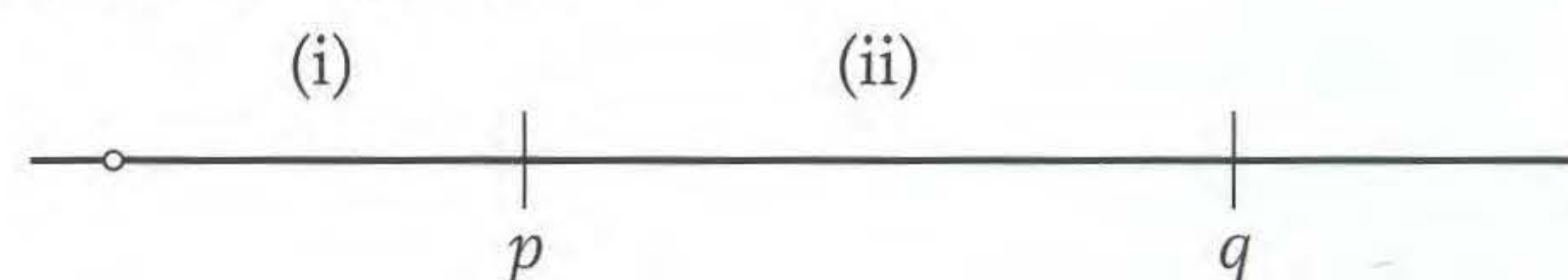
$$+ + + / + + + \times w s e / w s e$$

e a F_1 for submetida a um cruzamento-teste com $w s e$, e se for suposto que não há interferência nessa região do cromossomo, que proporção da prole será dos genótipos:

- | | |
|------------|------------|
| a. $+ + +$ | e. $+ + e$ |
| b. $w s e$ | f. $w s +$ |
| c. $+ s e$ | g. $w + e$ |
| d. $w + +$ | h. $+ s +$ |

57. Toda noite de sexta-feira, a estudante de genética Jean Allele, cansada de seus estudos, vai para a pista de boliche dos estudantes para relaxar. Mas, mesmo aí, ela é perseguida por seus estudos genéticos. A modesta pista de boliche tem apenas 4 bolas: duas vermelhas e duas azuis. Elas são lançadas para os pinos e, então, coletadas e retornadas em uma ordem aleatória, parando no final. À medida que o tempo passa, Jean nota padrões familiares das quatro bolas conforme elas voltam e param. Compulsivamente, ela conta os padrões diferentes. Que padrões ela vê, quais são suas frequências e qual a relevância desse assunto para a Genética?

58. Na análise de tétrades, o arranjo de ligação dos loci *p* e *q* são os seguintes:



Suponha que

- na região i, não há crossing em 88% das meioses e há um único crossing em 12% das meioses;
- na região ii, não há crossing em 80% das meioses e há um único crossing em 20% das meioses; e

• não há interferência (em outras palavras, a situação em uma região não afeta o que está ocorrendo na outra região).

Que proporções de tétrades serão dos seguintes tipos? **(a)** $M_I M_I$, DP; **(b)** $M_I M_I$, DNP; **(c)** $M_I M_{II}$, T; **(d)** $M_{II} M_I$, T; **(e)** $M_{II} M_{II}$, DP; **(f)** $M_{II} M_{II}$, DNP; **(g)** $M_{II} M_{II}$, T. **(Nota:** Aqui o padrão M escrito primeiro é o que pertence ao locus p .) **Pista:** O modo mais fácil de resolver o problema é começar calculando as frequências de ascas com crossings em ambas as regiões, região i, região ii e nenhuma região. Então, determine que padrões M_I e M_{II} resultam.

59. Para um experimento com uma levedura haplóide, você tem duas culturas diferentes. Cada uma crescerá em meio mínimo ao qual foi adicionado arginina, mas nenhum crescerá apenas em meio mínimo. (Meio mínimo é de sais inorgânicos mais açúcar.) Usando métodos apropriados, você induz as duas culturas a se reproduzirem. As células diplóides, então, dividem-se meioticamente e formam tétrades desordenadas. Alguns dos ascosporos crescerão em meio mínimo. Você classifica um grande número dessas tétrades quanto aos fenótipos ARG^- (requer arginina) e ARG^+ (independe de arginina) e registra os seguintes dados:

Segregação de ARG^- : ARG^+	Frequência (%)
4:0	40
3:1	20
2:2	40

a. Usando símbolos de sua escolha, atribua genótipos às duas culturas parentais. Para cada um dos três tipos de segregação, atribua genótipos aos segregantes.
b. Se houver mais de um locus controlando a necessidade de arginina, esses loci estão ligados?

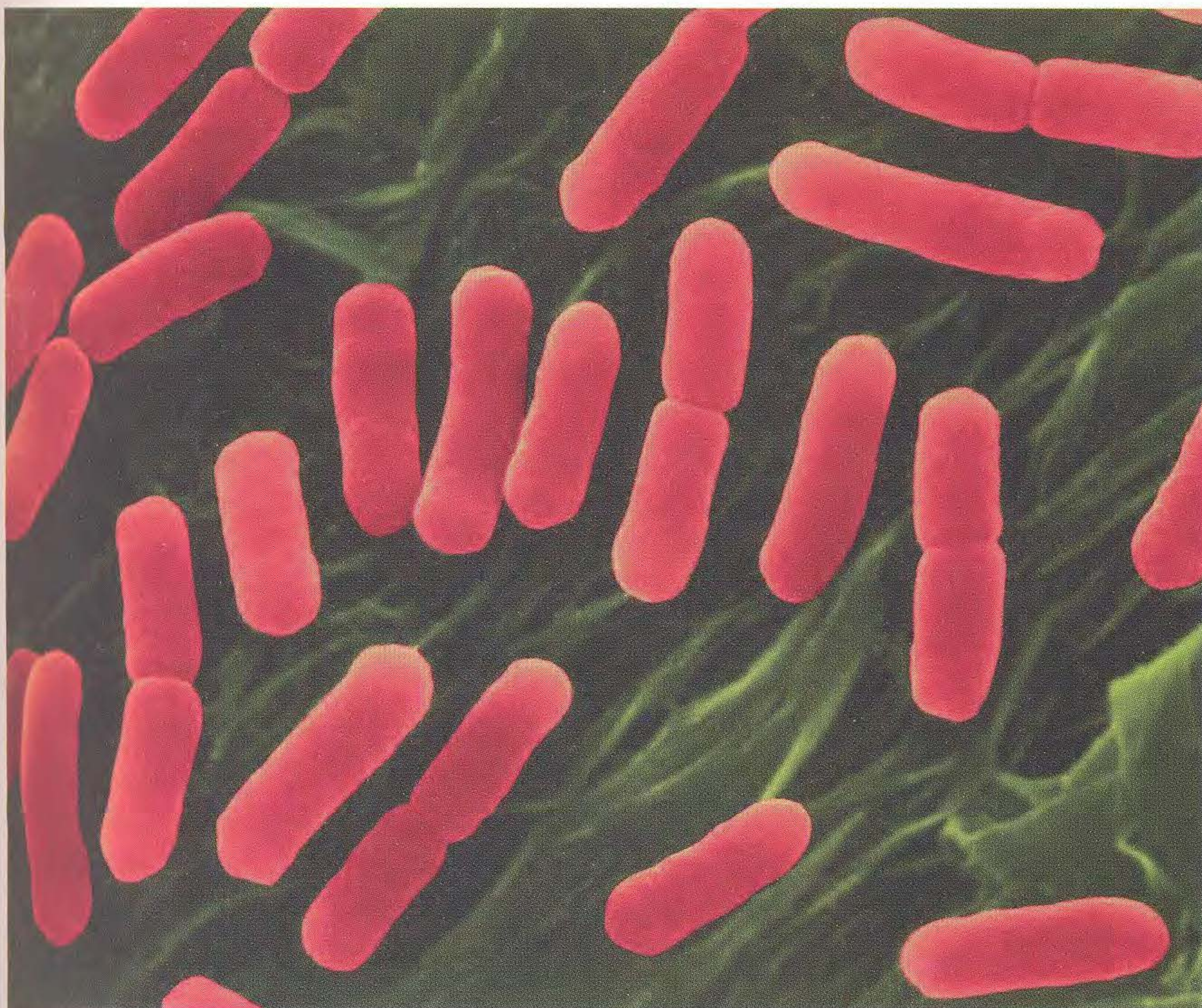
60. Uma análise de RFLP de duas linhagens puras $A/A \cdot B/B$ e $a/a \cdot b/b$ mostrou que a primeira era homozigota para um longo alelo de RFLP (l) e a segunda para um alelo curto (s). As duas foram cruzadas para formar uma F_1 , que foi então retrocruzada com a segunda linhagem pura. Uma prole de mil foi registrada do seguinte modo:

$Aa Bb ss$	9
$Aa Bb ls$	362
$aa bb ls$	11
$aa bb ss$	358
$Aa bb ss$	43
$Aa bb ls$	93
$aa Bb ls$	37
$aa Bb ss$	87

a. O que esses resultados nos dizem sobre ligação?
b. Desenhe um mapa se julgar apropriado.
c. Incorpore os fragmentos de RFLP ao seu mapa do modo da Figura 4.15.

A Genética de Bactérias e seus Vírus

5



Bactérias em divisão. [Dennis Kunkel Microscopy, Inc.]

A tecnologia do DNA é responsável pelos rápidos avanços que estão sendo feitos na genética de todos os organismos-modelo. É também um tópico de considerável interesse do domínio público. Os exemplos são os anúncios muito publicados das seqüências genômicas completas dos humanos e chimpanzés em anos recentes, e a popularidade da análise forense baseada no DNA em *shows* de televisão e filmes (Figura 5.1). Esses marcantes resultados, sejam em humanos, peixes, insetos, plantas ou fungos, são todos baseados no uso de tecnologias que permitem que pequenos pedaços de DNA sejam isolados, levados de célula a célula e amplificados em grandes amostras. Os sistemas sofisticados que permitem essas manipulações do DNA de qualquer organismo são quase todos derivados de bactérias e seus vírus. Assim, o avanço da Genética Moderna até seu presente estado de compreensão foi totalmente dependente do desenvolvimento da Genética de Bactérias, o tópico deste capítulo.

Muito embora a Genética de Bactérias tenha tornado possível a Genética Molecular, esta nunca foi a meta das pesquisas daquela. As bactérias são biologicamente importantes por si mesmas. São os organismos mais numerosos em nosso planeta. Contribuem para reciclar nutrientes tais como nitrogênio,

Perguntas Fundamentais

- Por que processos as bactérias trocam genes?
- Esses processos de troca podem ser usados para produzir fenótipos mutantes?
- Como os genomas de fagos interagem com genomas bacterianos?
- Como os genomas de fagos podem ser mapeados?

Tópicos

- 5.1 Trabalhando com microrganismos
- 5.2 Conjugação bacteriana
- 5.3 Transformação bacteriana
- 5.4 Genética de bacteriófagos
- 5.5 Transdução
- 5.6 Mapas físicos e mapas de ligação comparados

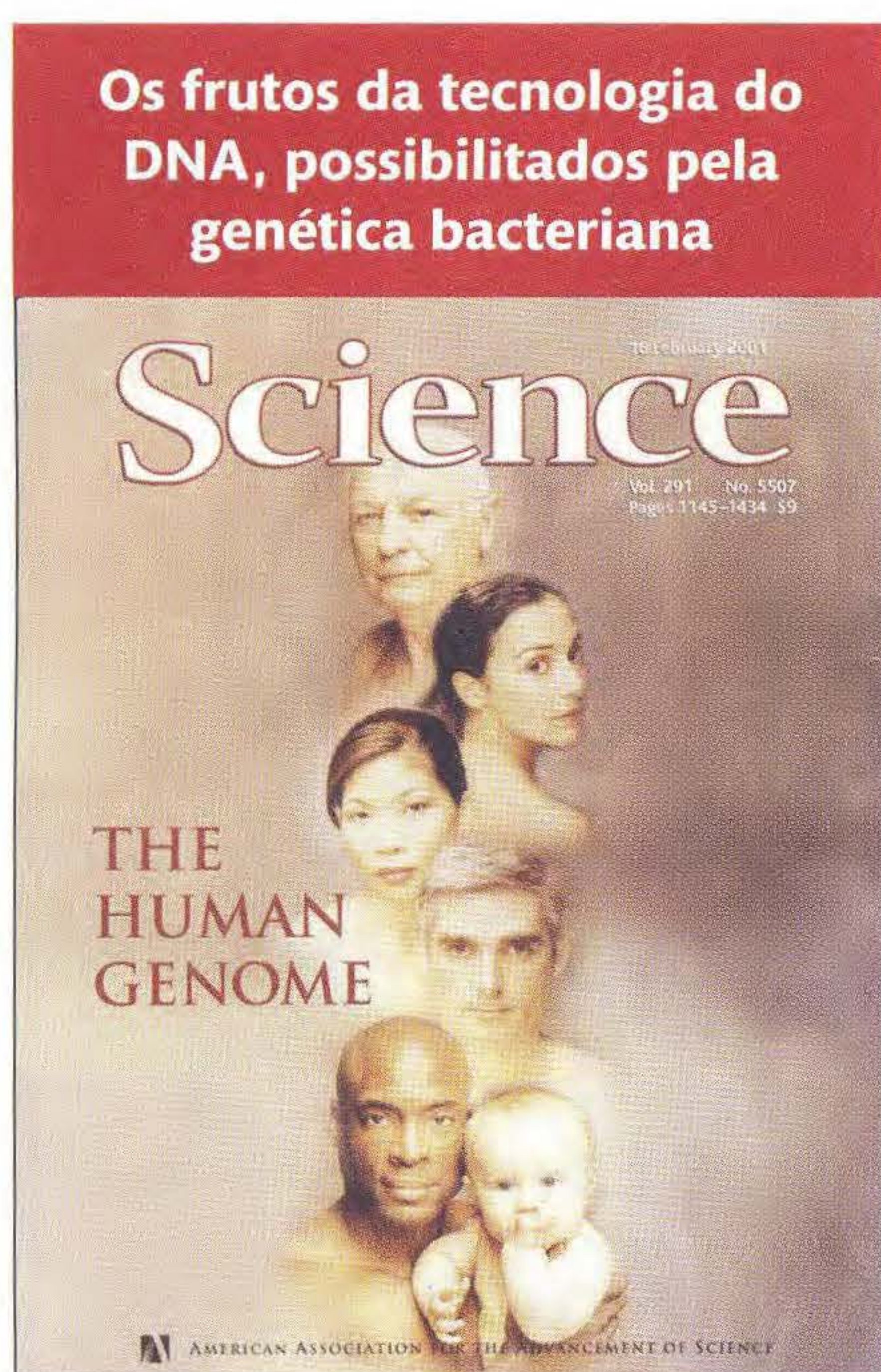


FIG. 5.1 Os marcantes resultados da moderna tecnologia do DNA, assim como o seqüenciamento do genoma humano, foram possíveis apenas porque a genética de bactérias levou à invenção de eficientes vetores de manipulação do DNA. [De *Science*, vol. 291, nº 5507 (February 16, 2001), pp. 1145-1434. Imagem de Ann E. Cutting. Reproduzida de AAAS, com permissão.]

enxofre e carbono nos ecossistemas. Algumas são agentes de doenças humanas, animais e de plantas. Outras vivem simbioticamente dentro de nossas bocas e intestinos. Além disso, muitos tipos de bactérias são úteis para a síntese industrial de uma ampla gama de produtos orgânicos. Assim, o ímpeto da dissecação genética das bactérias tem sido o mesmo que de organismos multicelulares, para compreender seu funcionamento biológico.

As bactérias pertencem a uma classe de organismos conhecidos como **procariontes**, que também incluem as algas cianofíceas (hoje classificadas como *cianobactérias*). Uma característica fundamental para definir os procariontes é que seu DNA não é circundado dentro de um núcleo delimitado por uma membrana. Como outros organismos, as bactérias têm genes compostos de DNA disposto em uma longa série de um "cromossomo". Entretanto, a organização de seu material genético é única sob vários aspectos. O genoma da maioria das bactérias é uma molécula única de DNA bifilamentar sob a forma de um círculo fechado. Além disso, as bactérias na natureza em geral contêm elementos extras de DNA chamados de plasmídeos. A maioria dos plasmídeos também são círculos de DNA, mas são muito menores que o genoma bacteriano principal.

As bactérias podem ser parasitadas por **vírus** específicos, chamados **bacteriófagos** ou, simplesmente, **fagos**. Os fagos e outros vírus são muito diferentes dos organismos que foram estudados até agora. Os vírus têm algumas propriedades em comum com organismos; por exemplo, seu material genético pode ser DNA ou RNA, constituindo um curto "cromossomo". Entretanto, a maioria dos biólogos encaram os vírus

como não-vivos porque eles não podem reproduzir-se sozinhos. Para se reproduzirem, devem parasitar células vivas e usar a maquinaria molecular dessas células. Assim, para estudar sua genética, eles devem se propagar nas células de seus organismos hospedeiros.

Quando os cientistas começaram a estudar bactérias e fagos, eles estavam naturalmente curiosos sobre seus sistemas hereditários. Obviamente, as bactérias e fagos devem ter sistemas hereditários, pois apresentam aspecto e função constantes de uma geração para a seguinte (são de tipo constante). Mas como funcionam esses sistemas hereditários? As bactérias, como os organismos eucarióticos unicelulares, reproduzem-se assexualmente por crescimento celular e divisão: uma célula torna-se duas. Essa reprodução assexuada é muito fácil de ser representada experimentalmente. Entretanto, existe alguma união de tipos diferentes para fins de reprodução sexual? Além disso, como os fagos bem menores se reproduzem? Eles se unem para um ciclo tipo sexual? Essas perguntas são estudadas neste capítulo.

Veremos que há uma variedade de processos hereditários em bactérias e fagos. Esses processos são curiosos devido à biologia básica dessas formas, mas elas também atuam como modelos, como fontes de percepção de processos genéticos que funcionam em *todos* os organismos. Para um geneticista, a atração dessas formas é que elas podem ser cultivadas em grandes números porque são muito pequenas. Conseqüentemente, é possível detectar e estudar *eventos genéticos raros* que são difíceis ou impossíveis de estudar em eucariontes.

Que processos hereditários são observados em procariontes? Na divisão celular assexual, o DNA é replicado, mas a repartição de novas cópias nas células-filhas é feita por um mecanismo bem diferente da mitose. Surgem mutantes? De fato, o processo de mutação ocorre nas células assexuadas do mesmo modo que ocorre em eucariontes, permitindo assim a dissecação genética da função bacteriana.

E quanto à reprodução sexual e à recombinação? Como as células e seus cromossomos são muito pequenos, os possíveis eventos tipo sexual são difíceis de observar, mesmo com um microscópio. Portanto, o enfoque geral para detectar fusão tipo sexual tem sido o genético, baseado na detecção de recombinantes. A lógica é que, se genomas diferentes vão juntos para a mesma célula, eles devem ocasionalmente produzir recombinantes. Contrariamente, se os recombinantes são detectados, com um marcador A de um genitor e B do outro, deve ter havido algum tipo de união "sexual". Assim, muito embora as bactérias e fagos não sofram meiose, o enfoque para a análise genética dessas formas é surpreendentemente similar ao dos eucariontes. A oportunidade para recombinação genética em bactérias pode surgir de diferentes modos, mas, em todos os casos, duas moléculas de DNA são levadas juntas. Entretanto, uma diferença importante dos eucariontes é que, nas bactérias, raramente existem dois cromossomos completos levados juntos. Geralmente, a união é de um cromossomo completo mais um fragmento do outro. As possibilidades são mostradas na Figura 5.2.

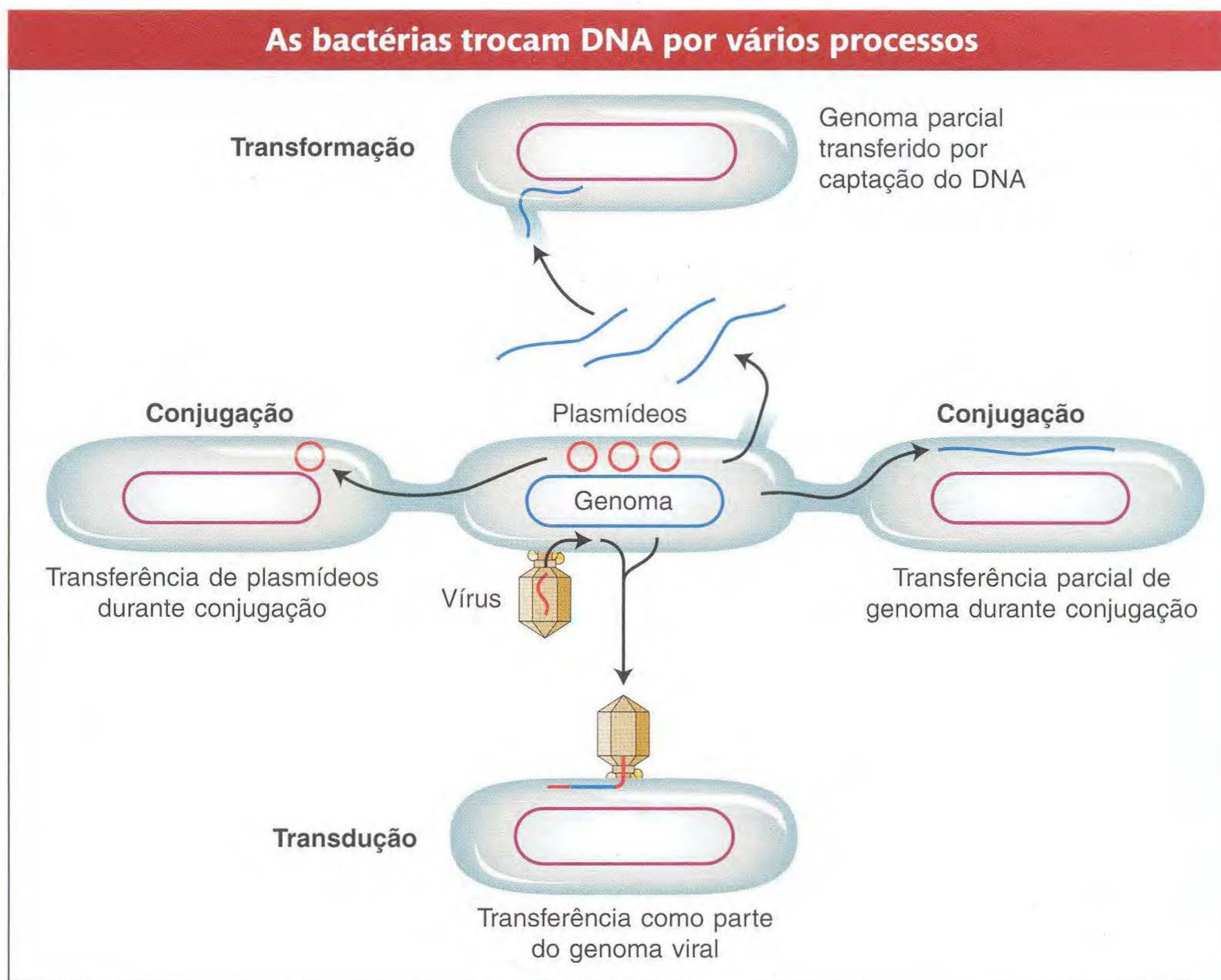


FIG. 5.2 O DNA bacteriano pode ser transferido de célula a célula de quatro modos: conjugação com transferência de plasmídeo, conjugação com transferência genômica parcial, transformação e transdução.

O primeiro processo de troca de genes a ser examinado no capítulo é a *conjugação*, que é o contato e fusão de duas células diferentes. Após a fusão, uma célula, chamada de doadora, às vezes transfere DNA de uma direção para outra célula. O DNA transferido pode ser parte do genoma bacteriano ou (raramente) todo esse. Em alguns casos, um dos elementos de DNA genômico chamado plasmídeo, se presente, é transferido. Qualquer fragmento genômico transferido pode se recombinar com o cromossomo receptor após a entrada.

Uma célula bacteriana também pode captar um pedaço de DNA do meio externo e incorporar esse DNA a seu próprio cromossomo, um processo chamado de *transformação*. Além disso, alguns fagos podem captar um pedaço de DNA de uma bactéria e injetá-lo em outra, onde ele pode ser incorporado ao cromossomo, em um processo chamado de *transdução*.

Os fagos podem sofrer recombinação quando dois genótipos diferentes infectam a mesma bactéria (**recombinação de fago**, não mostrada na Figura 5.2).

Antes de analisar esses modos de troca genética, consideremos os modos práticos de lidar com bactérias, que são muito diferentes dos usados em lidar com organismos multicelulares.

► 5.1 Trabalhando com Microrganismos

As bactérias dividem-se rapidamente e ocupam pouco espaço; logo, são convenientes para usar como organismos-modelo. Elas podem ser cultivadas em meio líquido ou em uma superfície sólida tal como gel de ágar, desde que os nutrientes básicos sejam fornecidos. Cada bactéria divide-se assexualmente em $1 \rightarrow 2 \rightarrow 4 \rightarrow 8 \rightarrow 16$ células, e assim por diante, até que os nutrientes sejam esgotados ou até que produtos tóxicos se acumulem a níveis que param o crescimento da população. Uma pequena quantidade de cultura líquida pode ser pipetada para uma placa de Petri contendo meio sólido de ágar e espalhar-se igualmente pela superfície com um espalhador estéril, em um processo chamado **plaqueamento** (Figura 5.3). As células dividem-se, mas, como não podem viajar pela superfície do gel, todas as células ficam juntas em um aglomerado. Quando essa massa atinge mais de 10^7 células, ela se torna visível a olho nu como uma **colônia**. Cada colônia distinta em uma placa foi derivada de uma única célula original. Os membros de uma colônia que têm um único ancestral genético são conhecidos como um **clone celular**.

Os mutantes bacterianos são fáceis de obter. Os mutantes nutricionais são um bom exemplo. As bactérias tipo selvagem são **prototróficas**, o que significa que elas podem crescer e

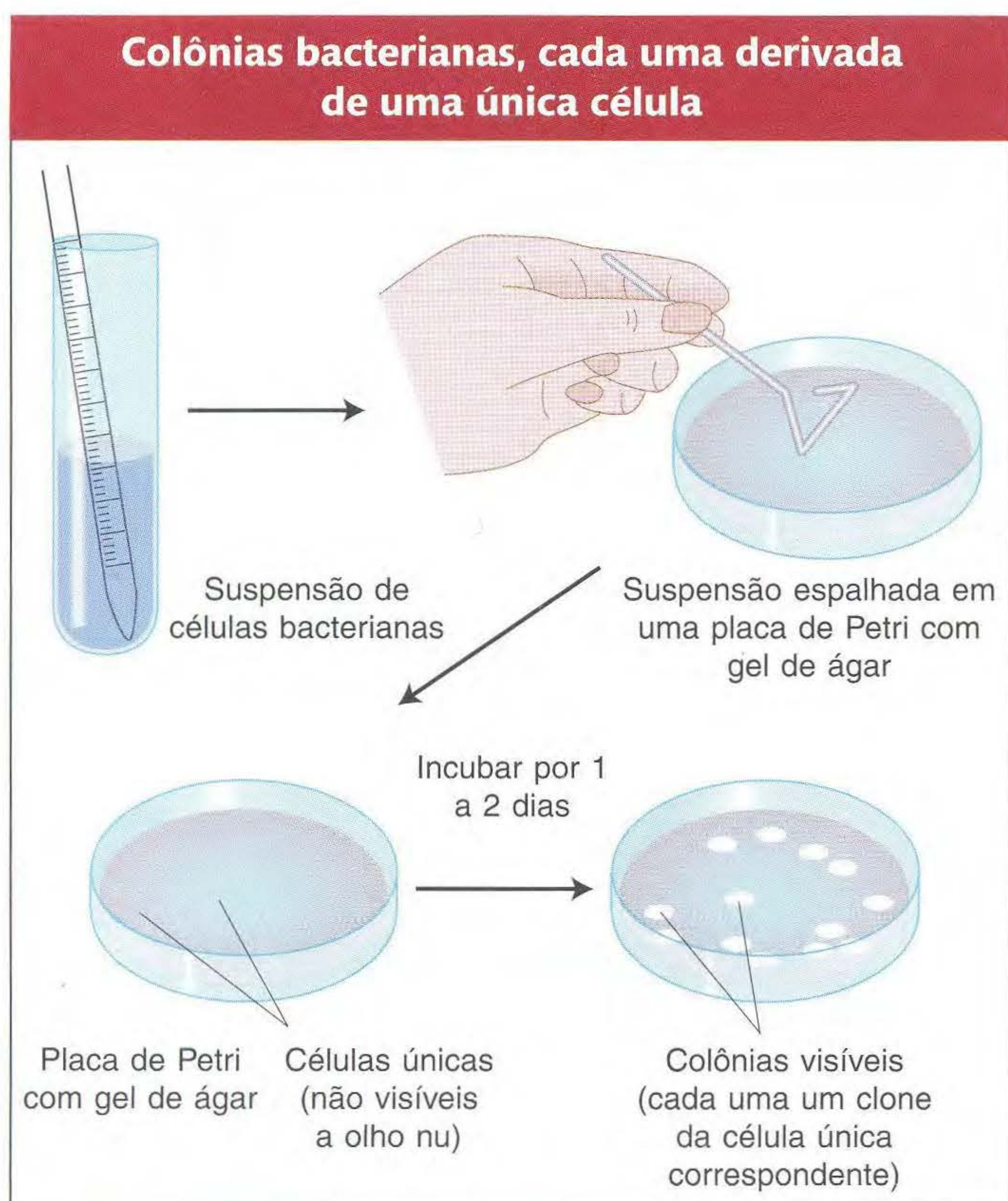


FIG. 5.3 Fenótipos bacterianos podem ser avaliados em suas colônias. Um estoque de células bacterianas pode ser cultivado em um meio líquido contendo nutrientes, e, então, um pequeno número de bactérias da suspensão líquida pode ser espalhada em meio sólido de ágar. Cada tipo de célula dará origem a uma colônia. Todas as células em uma colônia têm o mesmo genótipo e fenótipo.

se dividir em **meio mínimo**, um substrato contendo apenas sais inorgânicos, uma fonte de carbono para energia e água. De uma cultura prototrófica, podem ser obtidos mutantes **auxotróficos**: esses mutantes são células que não crescerão a menos que o meio contenha um ou mais blocos estruturais celulares específicos, tais como adenina, treonina ou biotina. Outro tipo de mutante útil difere do tipo selvagem na habilidade em usar uma fonte específica de energia; por exemplo, o tipo selvagem (lac^+) pode usar lactose e crescer, enquanto um mutante (lac^-) não pode. A Figura 5.4 mostra um outro



FIG. 5.4 Bactérias tipo selvagem capazes de usar lactose como fonte de energia (lac^+) coram-se em vermelho na presença desse corante indicador. As células não-coradas são mutantes incapazes de usar lactose (lac^-). [Jeffrey H. Miller.]

modo de distinguir colônias lac^+ e lac^- usando um corante. Em outra categoria de mutantes, enquanto os tipos selvagens são suscetíveis a um inibidor, tal como o antibiótico estreptomicina, os **mutantes resistentes** podem se dividir e formar colônias na presença de um inibidor. Todos esses tipos de mutantes permitem que o geneticista distinga linhagens individuais diferentes, fornecendo assim **marcadores genéticos** (alelos marcadores) para rastrear genomas e células em experimentos. O Quadro 5.1 resume alguns fenótipos bacterianos mutantes e seus símbolos genéticos.

As seções seguintes documentam a história de vários processos pelos quais os genomas bacterianos se recombina. Os métodos históricos são interessantes em si mesmos, mas também servem para introduzir os diversos processos de recombinação, bem como as técnicas analíticas que ainda são aplicáveis hoje em dia.

► 5.2 Conjugação Bacteriana

Os primeiros estudos de genética bacteriana revelaram o processo inesperado de conjugação bacteriana.

Quadro 5.1 Alguns Símbolos Genotípicos Usados em Genética de Bactérias

Símbolo	Característica ou fenótipo associado ao símbolo
bio^-	Requer adição de biotina como suplemento ao meio mínimo
arg^-	Requer arginina adicionada a um suplemento ao meio mínimo
met^-	Requer metionina adicionada a um suplemento ao meio mínimo
lac^-	Não usa lactose como fonte de carbono
gal^-	Não usa galactose como fonte de carbono
str^r	Resistente ao antibiótico estreptomicina
str^s	Sensível ao antibiótico estreptomicina

Nota: Meio mínimo é o meio sintético básico para crescimento de bactérias sem suplementos de nutrientes.

Descoberta da conjugação

As bactérias possuem processos similares à reprodução sexual e recombinação? A pergunta foi respondida pelo trabalho experimental elegante de Joshua Lederberg e Edward Tatum, que, em 1946, descobriram um processo tipo sexual no que se tornou o principal modelo para a genética de bactérias a *Escherichia coli* (veja o boxe Organismo-modelo). Eles estavam estudando duas linhagens de *E. coli* com conjuntos diferentes de mutações auxotróficas. A linhagem A só cresceria se o meio fosse suplementado com metionina e biotina; a linhagem B só cresceria se fosse suplementada com treonina, leucina e tiamina. Assim, podemos chamar as linhagens de

linhagem A: $met^- bio^- tre^+ leu^+ tia^+$

linhagem B: $met^+ bio^+ tre^- leu^- tia^-$

A Figura 5.5a mostra uma forma simplificada de representar seu experimento. As linhagens A e B foram misturadas, incubadas por um tempo e, então, plaqueadas em meio mínimo, no qual nenhum auxotrófico podia crescer. Uma pequena minoria de células (1 em 10^7) crescia como prototróficos e, assim, devem ter sido tipo selvagem, tendo recuperado a habilidade em crescer sem a ajuda de nutrientes. Algumas

das placas receberam apenas a linhagem A de bactérias e algumas só a linhagem B de bactérias para agir como controles, mas nenhum prototrófico surgiu desses plaqueamentos. A Figura 5.5b ilustra o experimento em mais detalhes. Esses resultados sugerem que alguma forma de recombinação de genes ocorreu entre os genomas das duas linhagens para produzir os prototróficos.

Poderia ser questionado que as células das duas linhagens de fato não trocam genes, mas sim vazam substâncias que outras células podem absorver e usar para crescer. Essa possibilidade de “alimentação cruzada” foi excluída por Bernard Davis do seguinte modo. Ele construiu um tubo em forma de U no qual os dois braços eram separados por um fino filtro. Os poros do filtro eram muito pequenos para permitir a passagem de bactérias, mas grandes o suficiente para permitir a fácil passagem de quaisquer substâncias dissolvidas (Figura 5.6). A linhagem A foi posta em um braço, e a linhagem B no outro. Após as linhagens terem sido incubadas por um tempo, Davis testou os conteúdos de cada braço para ver se havia alguma prototrófica, mas não foi encontrada nenhuma. Em outras palavras, o *contato físico* entre as duas linhagens era necessário para a formação de células tipo selvagem. Era como se algum tipo de união de genoma tivesse ocorrido, e



Organismo-modelo *Escherichia coli*

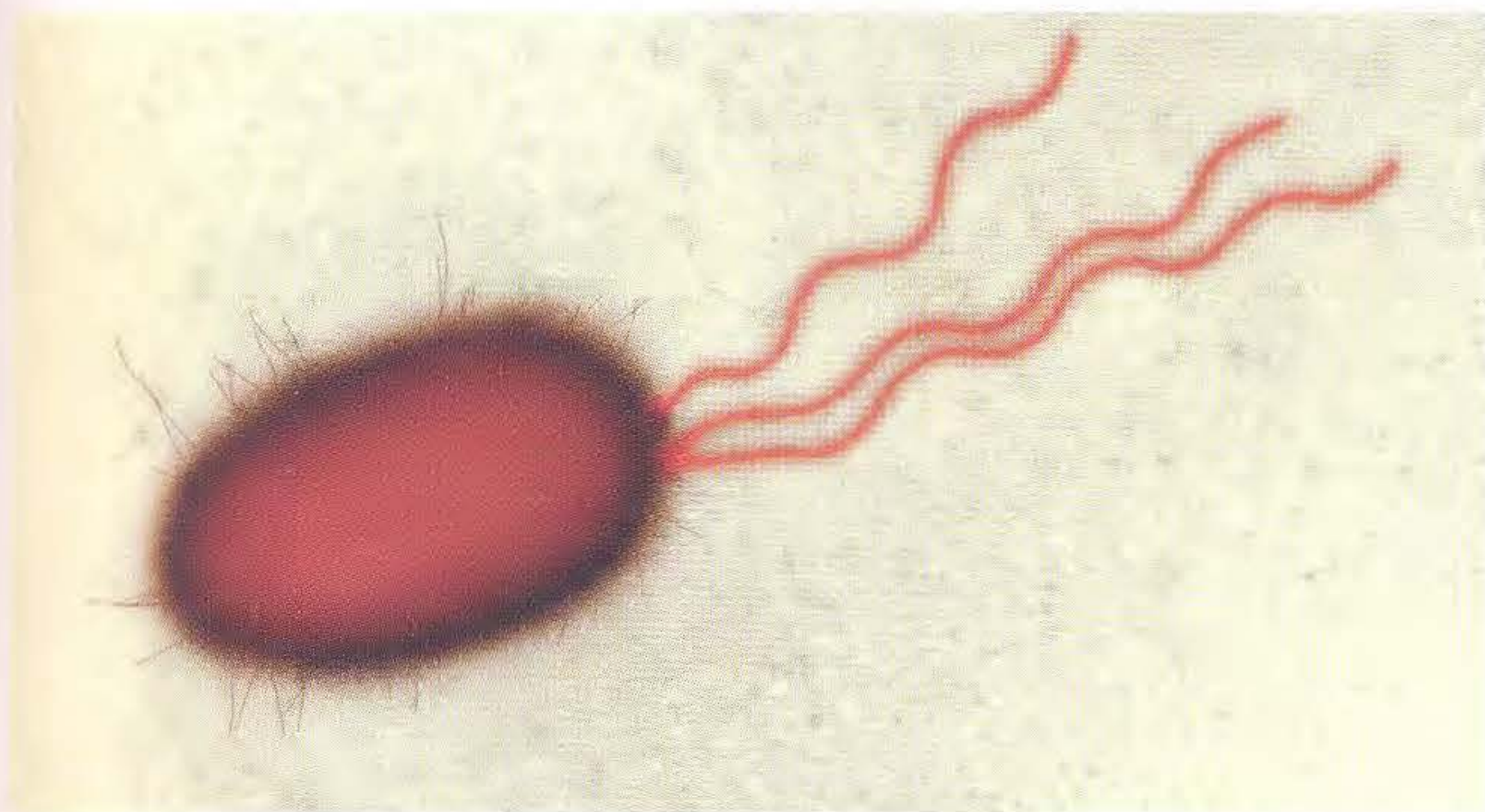
O microscopista do século XVII Antony van Leeuwenhoek foi provavelmente o primeiro a ver células bacterianas e reconhecer seu pequeno tamanho: “Existe mais vida na boca de um homem do que homens em todo o reino.” Entretanto, a Bacteriologia só começou de fato no século XIX. Nos anos de 1940, Joshua Lederberg e Edward Tatum fizeram a descoberta que lançou a Bacteriologia no campo borbulhante da Genética: eles descobriram que, em algumas bactérias, havia um tipo de ciclo sexual incluindo um processo de crossing. O organismo que eles escolheram para esse experimento tornou-se o modelo não só para a genética de procariontes, mas sim para toda a genética. O organismo foi a *Escherichia coli*, uma bactéria denominada

em homenagem a seu descobridor, o bacteriologista alemão do século XIX Theodore Escherich.

A escolha da *E. coli* foi feliz porque ficou provado que ela tem muitas características adequadas para a pesquisa genética, sendo uma delas facilmente obtida, pois vive no intestino de humanos e de outros animais. No nosso organismo, ela é um simbiote benigno, mas, ocasionalmente, causa infecções nas vias urinárias e diarreia.

A *E. coli* tem um único cromossomo circular de 4,6 Mb de tamanho. Dos seus 4.000 genes sem íntrons, cerca de 35% são de função desconhecida. O ciclo sexual é possibilitado pela ação de um plasmídeo extragenômico chamado F, que confere um tipo de “masculinidade”. Outros plasmídeos levam genes cujas funções equipam a célula para a vida em ambientes específicos, tais como genes resistentes a drogas. Esses plasmídeos foram adaptados como *vetores* gênicos, que são transportadores de genes que formam a base das transferências gênicas no centro da engenharia genética moderna.

E. coli é unicelular e cresce por simples divisão celular. Devido a seu pequeno tamanho ($\pm 1 \mu m$), *E. coli* pode ser cultivada em grandes números e sujeita a intensa seleção e triagem para raros eventos genéticos. A pesquisa de *E. coli* representa o começo do esclarecimento da “caixa preta” em genética: por meio da seleção e análise de mutantes, os trabalhos da maquinaria genética puderam ser deduzidos, muito embora fosse muito pequena para ser vista. Os fenótipos tais como tamanho da colônia, resistência a drogas, uso de fonte de carbono e produção de corantes tomaram o lugar dos fenótipos visíveis da genética dos eucariontes.



Micrografia eletrônica de uma *E. coli* mostrando flagelos longos, usados para locomoção, e fímbrias pelos proteináceos que são importantes para ancorar as células a tecidos animais. (Os pili sexuais não são mostrados nessa micrografia.) [Dr. Dennis Kunkel/Visuals Unlimited.]

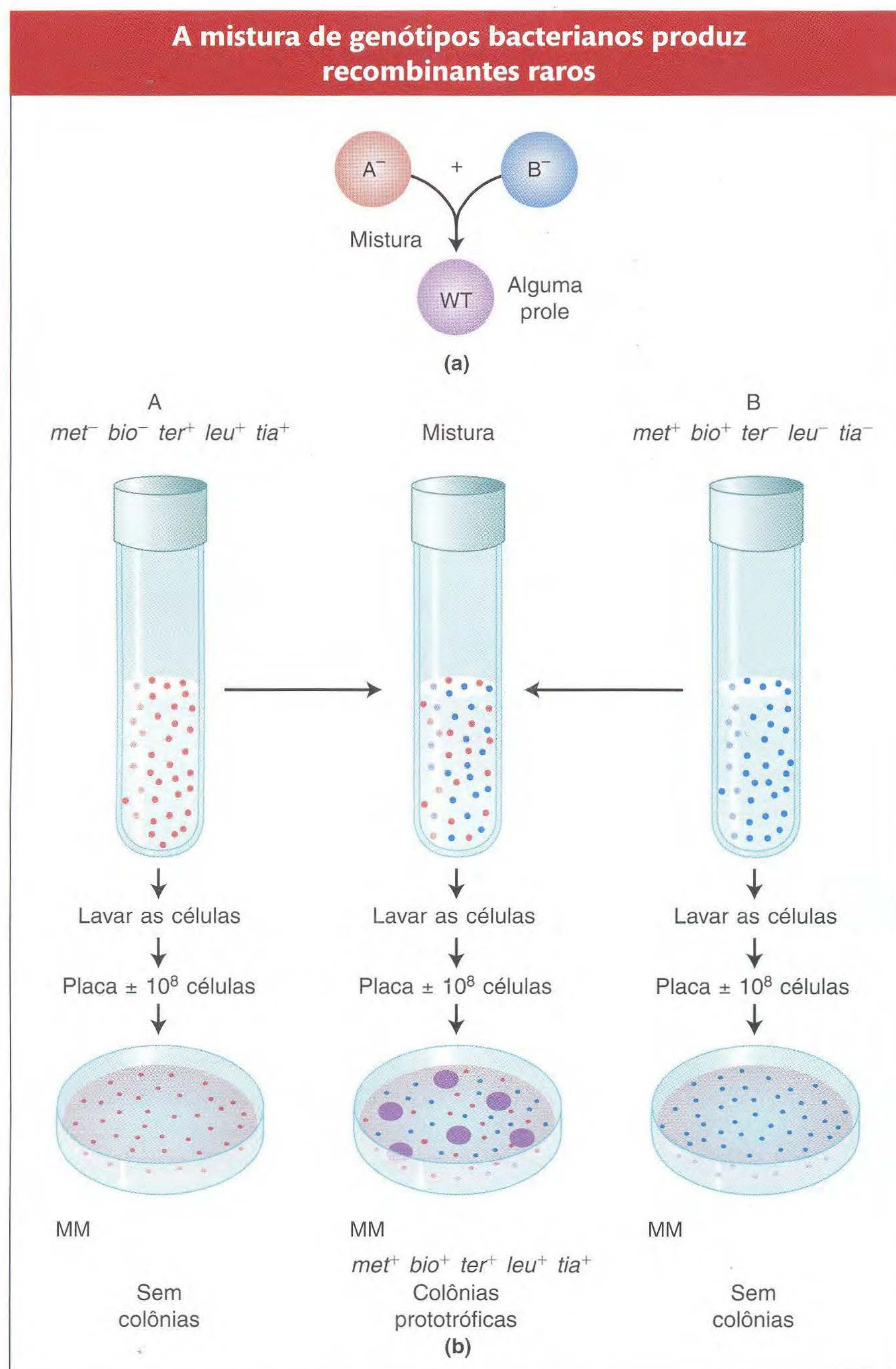


FIG. 5.5 Com o uso desse método, Lederberg e Tatum demonstraram que a recombinação genética entre genótipos bacterianos é possível. (a) O conceito básico: duas culturas auxotróficas (A^- e B^-) são misturadas, gerando tipos selvagens prototróficos (WT). (b) Células do tipo A ou tipo B não podem crescer em um meio não-suplementado (mínimo) (MM), porque A e B levam mutações que causam a incapacidade em sintetizar constituintes necessários para o crescimento celular. Quando A e B são misturados por algumas horas e então plaqueados, algumas colônias aparecem na placa de ágar. Essas colônias derivam de células únicas nas quais o material genético foi trocado; elas são, portanto, capazes de sintetizar todos os constituintes necessários ao metabolismo.

tinha sido produzida uma genuína recombinação. A união física das bactérias pode ser confirmada ao microscópio eletrônico, e hoje é chamada de **conjugação** (Figura 5.7).

Descoberta do fator fertilidade (F)

Em 1953, William Hayes descobriu que, nos tipos de “cruzamentos” descritos aqui, os genitores conjugantes agiram *desi-*

gualmente (mais adiante veremos os modos para demonstrar essa participação desigual). Um genitor (e *apenas* esse genitor) parecia transferir parte de seu genoma ou todo ele para outra célula. Assim, uma célula atua como uma **doadora**, e a outra célula atua como uma **receptora**. Esse “cruzamento” é muito diferente dos cruzamentos eucarióticos nos quais os genitores contribuem igualmente com genomas nucleares.

Não são produzidos recombinantes sem contato celular

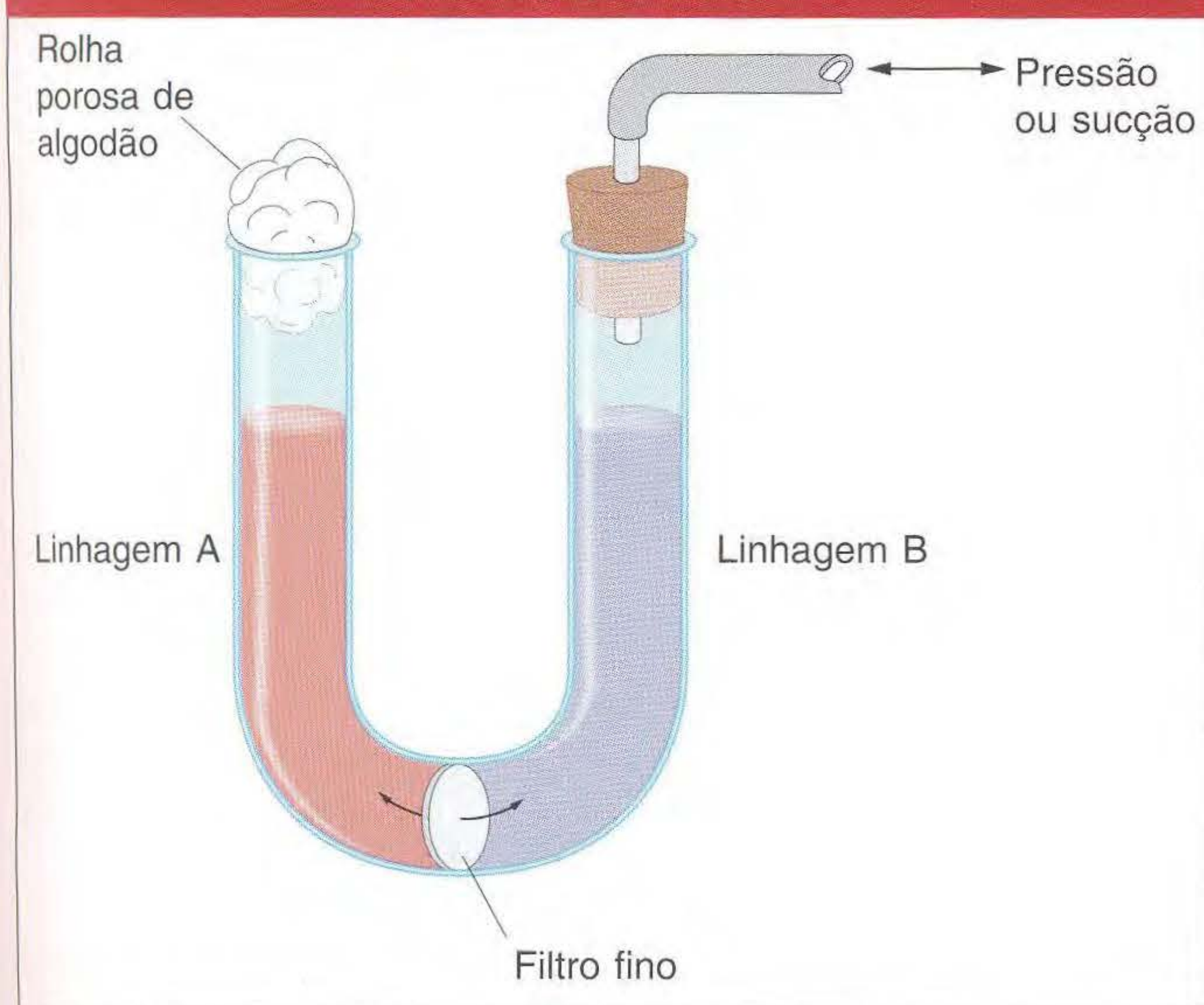


FIG. 5.6 As linhagens bacterianas auxotróficas A e B são cultivadas em ambos os lados do tubo em U. O líquido pode ser passado entre os braços aplicando pressão ou sucção, mas as células bacterianas não podem passar pelo filtro. Após incubação e plaqueamento, nenhuma colônia recombinante cresce no meio mínimo.

A transferência de material genético na conjugação de *E. coli* não é recíproca. Uma célula, a doadora, transfere parte de seu genoma para outra célula, que atua como receptora.

Por acaso, Hayes descobriu uma variante dessa linhagem doadora original que não produziria recombinantes no cruzamento com a linhagem receptora. Aparentemente, a linhagem tipo doadora tinha perdido a habilidade em transferir o material genético e tinha mudado para uma linhagem tipo

Conjugação bacteriana usando pili



FIG. 5.7 Uma célula doadora estende uma ou mais projeções, ou pili, que se ligam a uma célula receptora e ligam as duas bactérias. [Dennis Kunkel Microscopy, Inc.]

receptora. Ao trabalhar com essa variante doadora “estéril”, Hayes descobriu que ela podia recuperar a habilidade em agir como uma doadora por associação com outras linhagens doadoras. De fato, a habilidade de doadora era transmitida rápida e efetivamente entre as linhagens durante a conjugação. Um tipo de “transferência infecciosa” de algum fator parecia estar ocorrendo. Ele sugeriu que a habilidade de doadora é em si um estado hereditário, imposto por um **fator de fertilidade (F)**. As linhagens que possuem F podem doar e são designadas F^+ . As linhagens que não têm F não podem doar e são receptoras, chamadas de F^- .

Hoje sabemos muito mais sobre F. Ele é um exemplo de uma pequena molécula de DNA circular não-essencial, chamada de **plasmídeo**, que pode replicar-se no citoplasma independentemente do cromossomo hospedeiro. A Figura 5.8 mostra como as bactérias podem transferir plasmídeos tais como F. O plasmídeo F dirige a síntese dos pili (singular, pilus), projeções que iniciam o contato com uma receptora (veja Figuras 5.7 e 5.8) e a coloca mais perto. O DNA F na célula doadora faz uma cópia unifilamentar de si mesmo por um mecanismo peculiar chamado de **replicação em círculo rolante**. O plasmídeo circular “rola”, e, à medida que gira, ele desenrola uma cópia unifilamentar, como uma linha de pescar. Essa cópia passa por um poro para a célula receptora, onde o outro filamento é sintetizado, formando uma dupla hélice. Assim, uma cópia de F fica na doadora e a outra aparece na receptora, como mostrado na Figura 5.8. Note que o genoma de *E. coli* é representado como um único cromossomo circular na Figura 5.8. (Mais adiante examinaremos a evidência disso.) A maioria dos genomas bacterianos são circulares, uma característica bem diferente dos cromossomos nucleares eucarióticos. Veremos que essa característica leva a muitas idiosincrasias da genética bacteriana.

Linhagens Hfr

Uma novidade importante surgiu quando Luca Cavalli-Sforza descobriu um derivado de uma linhagem F^+ com duas propriedades incomuns:

1. Ao ser cruzada com linhagens F^- , essa nova linhagem produziu 1.000 vezes mais recombinantes que a linhagem F^+ normal. Cavalli-Sforza chamou essa derivada de linhagem **Hfr** para representar sua habilidade em promover uma *high frequency*, alta frequência, de recombinação.
2. Nos cruzamentos $Hfr \times F^-$, praticamente nenhum dos genitores F^- foi convertido em F^+ ou em Hfr. Esse resultado contrasta com os cruzamentos $F^+ \times F^-$, nos quais, como vimos, a transferência infecciosa de F resulta em uma grande proporção de genitores F^- sendo convertidos em F^+ .

Ficou aparente que uma linhagem Hfr resulta da integração do fator F ao cromossomo, como representado na Figura 5.9. Agora podemos explicar a primeira propriedade incomum das linhagens Hfr. Durante a conjugação, o fator F inserido no cromossomo ativa eficientemente parte desse cromossomo ou todo ele na célula F^- . O fragmento cromos-

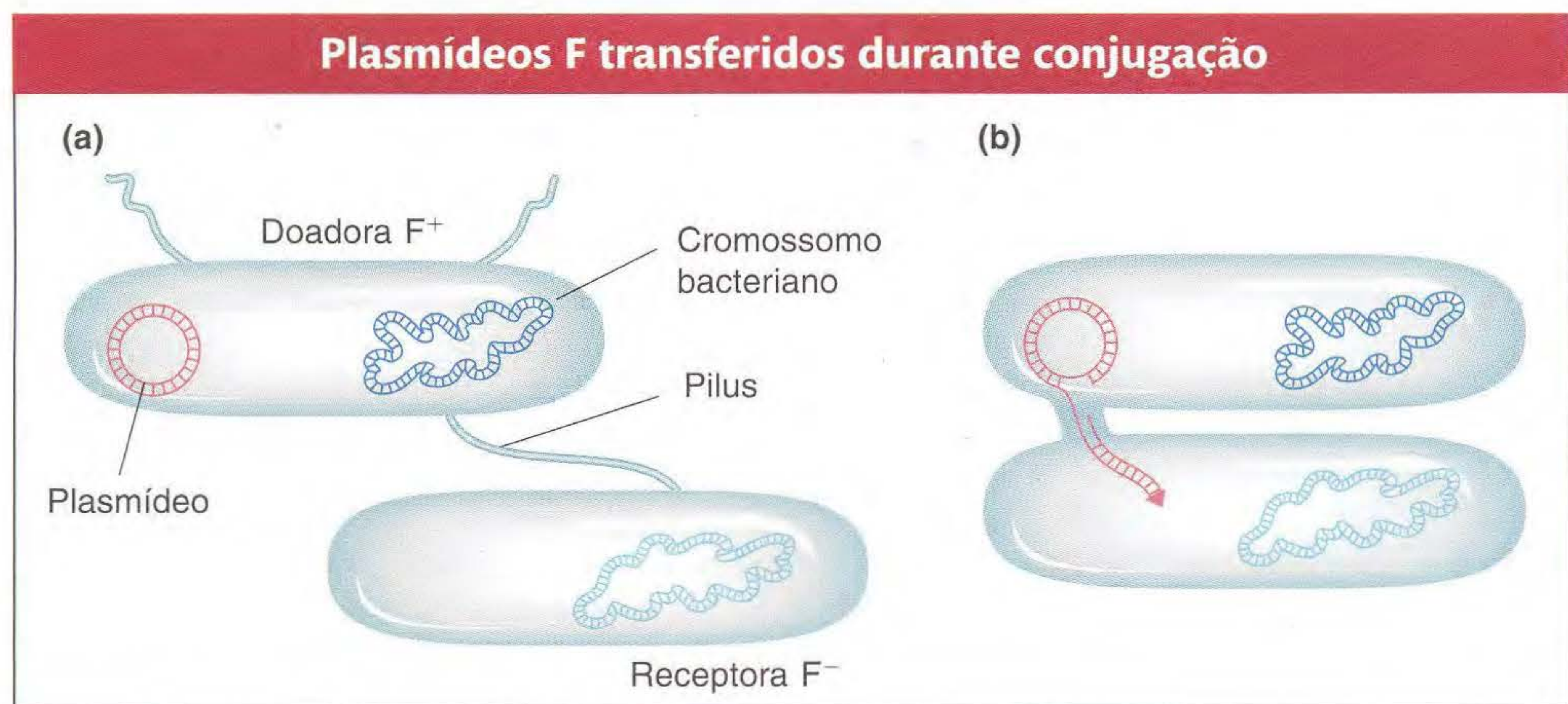


FIG. 5.8 (a) Durante a conjugação, o pilus junta as duas bactérias. (b) Em seguida, uma ponte (essencialmente um poro) forma-se entre as duas células. Uma cópia unifilamentar do DNA do plasmídeo é produzida na célula doadora e, então, passa para a bactéria receptora, onde o filamento único, servindo como molde, é convertido na dupla hélice.

sômico pode então participar da recombinação com o cromossomo receptor. Os raros recombinantes observados por Lederberg e Tatum nos cruzamentos $F^+ \times F^-$ eram devidos à espontânea, mas rara, formação de células Hfr na cultura F^+ . Cavalli-Sforza isolou exemplos dessas raras células de culturas F^+ e observou que, de fato, elas agora agiam como verdadeiras Hfr.

Uma Hfr morre após doar seu material cromossômico a uma F^- ? A resposta é não. Como o plasmídeo F, o cromossomo Hfr replica-se e transfere um único filamento para a célula F^- durante a conjugação. Que o DNA transferido é um filamento único pode ser demonstrado visualmente com o uso de corantes especiais e anticorpos, como mostrado na Figura 5.10. A replicação do cromossomo garante um cromossomo completo para a célula doadora após a reprodução. O filamento transferido é convertido em uma dupla hélice na célula receptora, e os genes doadores podem tornar-se incorporados ao cromossomo receptor por crossings, criando uma célula recombinante (Figura 5.11). Se não houver recombi-

nação, os fragmentos de DNA transferidos são simplesmente perdidos no curso da divisão celular.

Transmissão linear dos genes Hfr a partir de um ponto fixo Uma visão mais clara do comportamento das linhagens

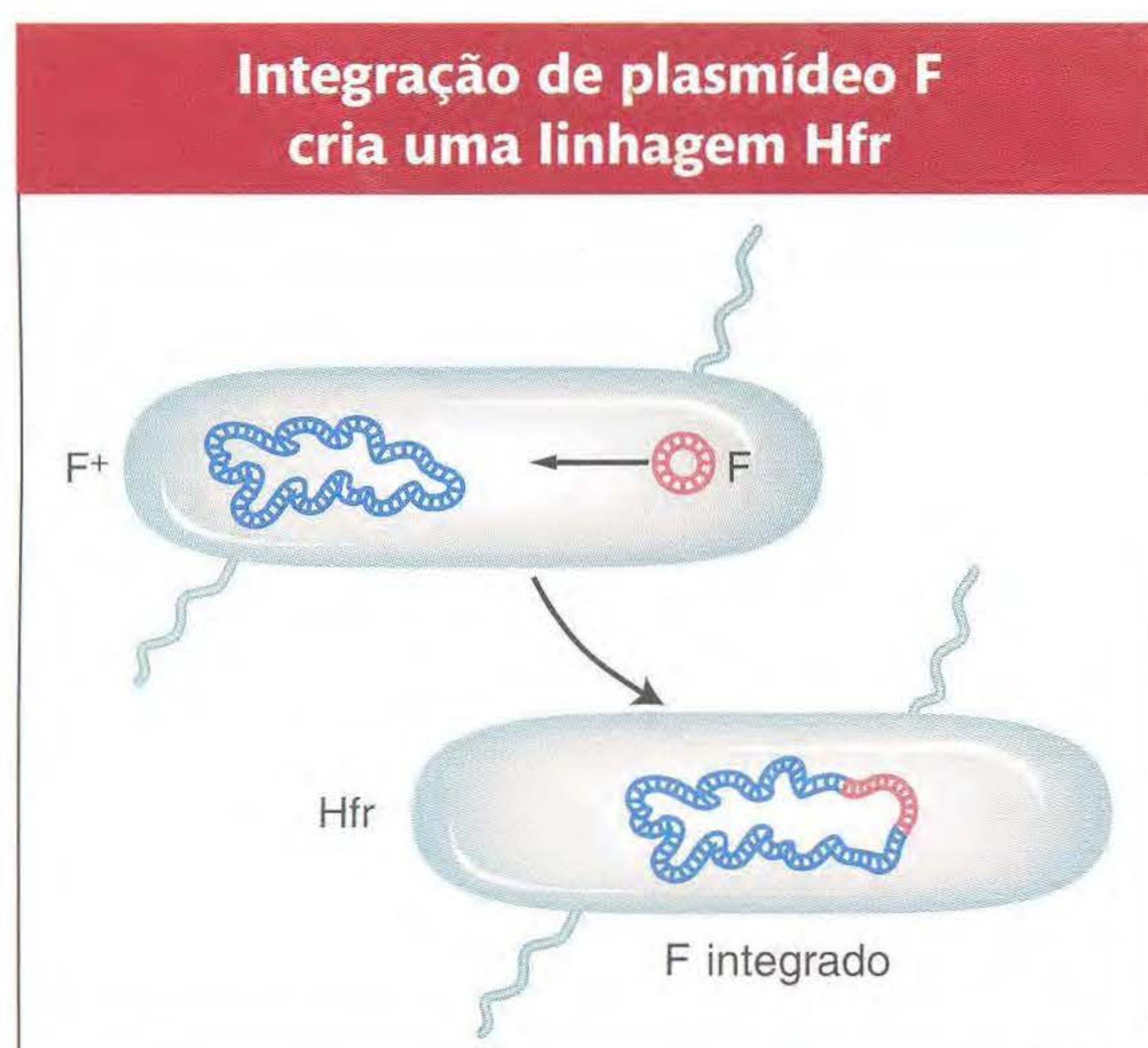


FIG. 5.9 Em uma linhagem F^+ , o plasmídeo F livre ocasionalmente se integra ao cromossomo de *E. coli*, criando uma linhagem Hfr.

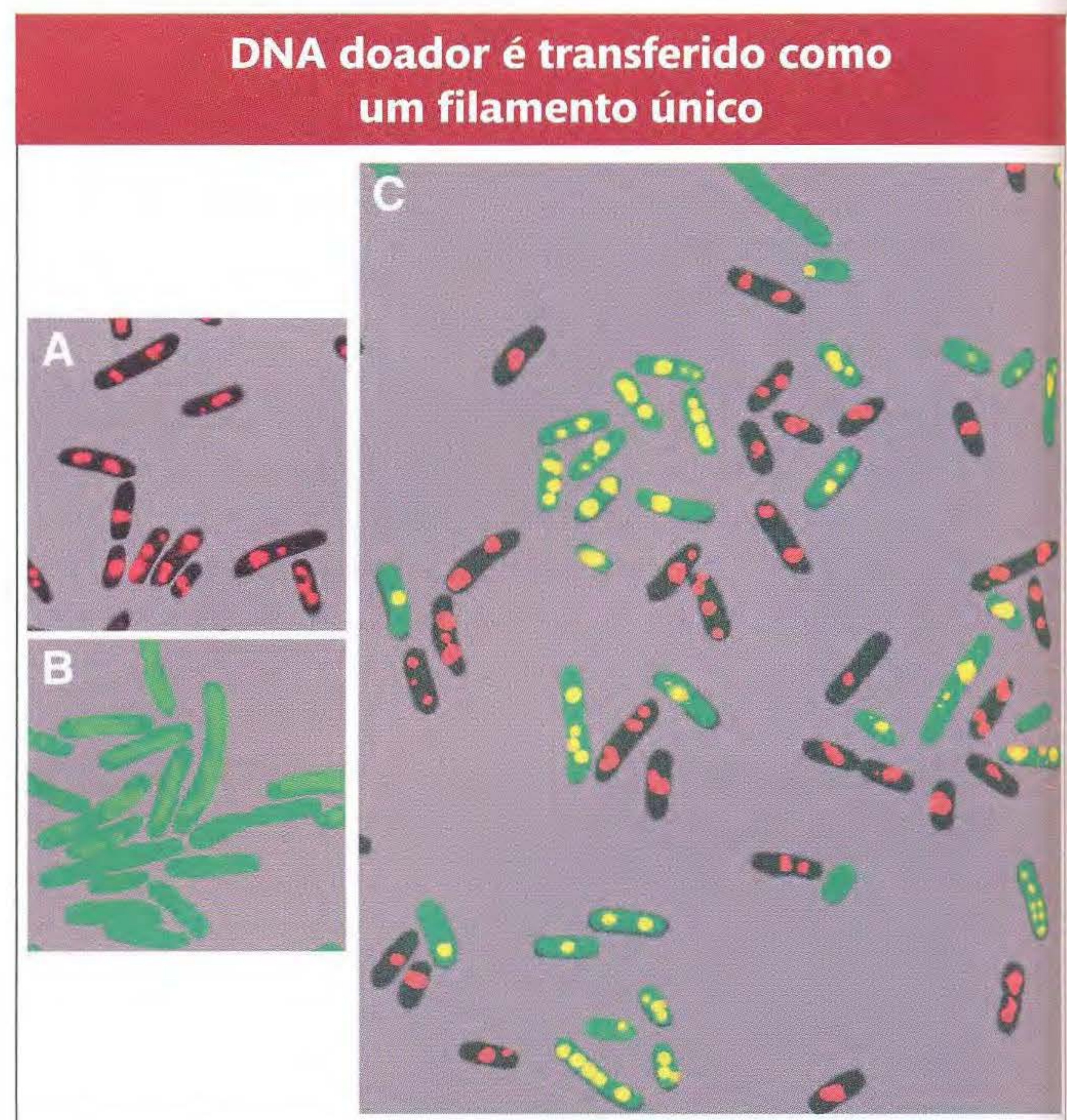


FIG. 5.10 As fotos mostram a visualização de DNA unifilamentar transferido na conjugação de *E. coli*, com o uso de anticorpos fluorescentes especiais. As linhagens Hfr parentais (A) são pretas com DNA vermelho. O vermelho é da ligação de um anticorpo com uma proteína normalmente ligada ao DNA. As células receptoras F^- (B) são verdes devido à presença do gene para uma proteína de água viva que fluoresce em verde; e, como são mutantes para certo gene, a proteína de seu DNA não se liga a anticorpo. Quando o DNA unifilamentar entra na receptora, ela promove uma ligação atípica dessa proteína, que fluoresce em amarelo nesse fundo. A Parte C mostra Hfr (inalterada) e exconjugantes (células que sofreram conjugação) com DNA transferido em amarelo. Algumas células F^- não-reproduzidas são visíveis. [De M. Kohiyama, S. Hiraga, I. Matic and M. Radman, "Bacterial Sex: Playing Voyeurs 50 Years Later." *Science* 301, 2003, p. 803, Fig. 1. Reproduzida de AAAS, com permissão.]

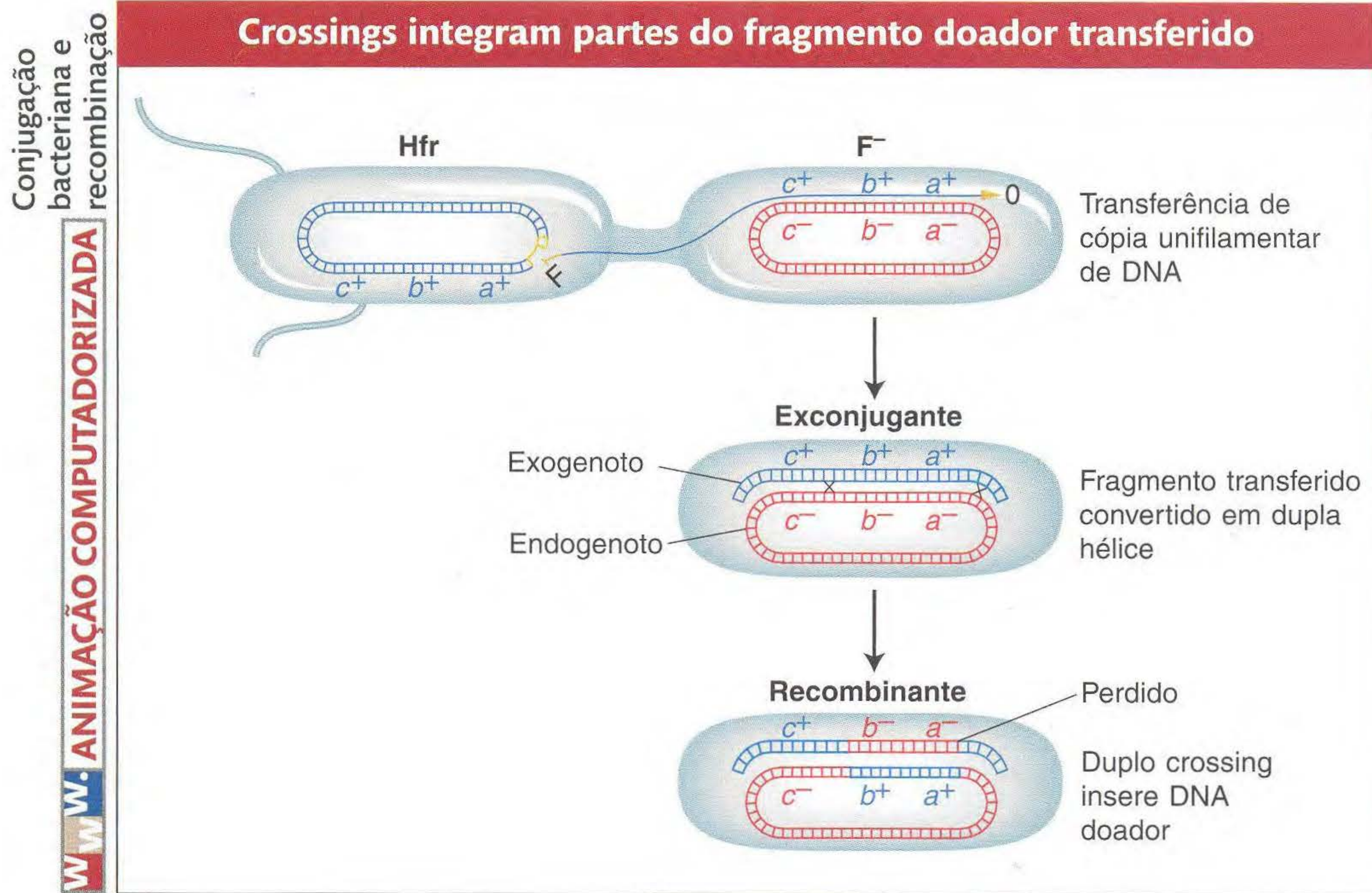


FIG. 5.11 Após a conjugação, são necessários crossings para integrar genes do fragmento doador ao cromossomo receptor e, assim, tornar-se uma parte estável de seu genoma.



Hfr foi obtido em 1957, quando elie Wollman e François Jacob investigaram o padrão de transmissão de genes Hfr para células F⁻ durante um cruzamento. Eles cruzaram

$$\text{Hfr } azi^r \text{ ton}^r \text{ lac}^+ \text{ gal}^+ \text{ str}^s \times \text{F}^- \text{ azi}^s \text{ ton}^s \text{ lac}^- \text{ gal}^- \text{ str}^r$$

(Os expoentes “r” e “s” significam resistente e sensível, respectivamente.) Em épocas específicas após a mistura, eles obtiveram amostras, as quais foram cada uma colocada em um liquidificador por alguns segundos para separar os pares de células reprodutoras. Esse processo é chamado de **reprodução interrompida**. A amostra foi então plaqueada em um meio contendo estreptomicina para matar as células doadoras Hfr, que tinham o alelo de sensibilidade *str*^s. As células sobreviventes *str*^r foram então testadas quanto à presença de alelos do genoma doador. Qualquer célula *str*^r tendo um alelo doador deve ter tomado parte na conjugação. Tais células são chamadas de **exconjugantes**. Os resultados são plotados na Figura 5.12a, mostrando o curso de entrada de cada alelo doador *azi*^r, *ton*^r e *gal*⁺. A Figura 5.12b mostra a transferência dos alelos Hfr.

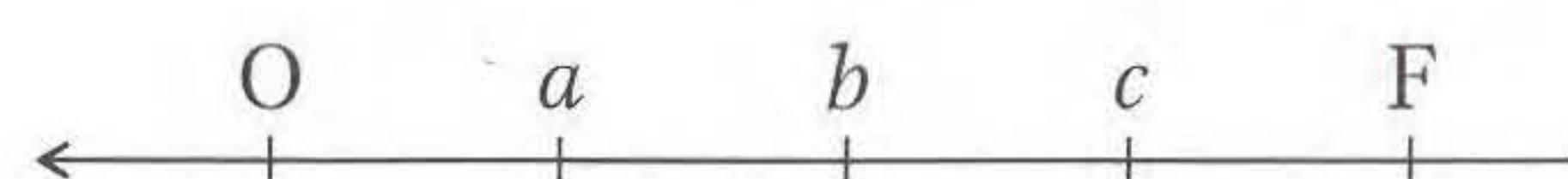
Os principais elementos nesses resultados são

1. Cada alelo doador primeiro aparece nas receptoras F⁻ em um momento específico após começar a reprodução.
2. Os alelos doadores aparecem em uma seqüência específica.
3. Os alelos doadores posteriores estão presentes em menos células receptoras.

Juntando todas essas observações, Wollman e Jacob deduziram que, na Hfr conjugante, a transferência do DNA unifi-

lamentar começa a partir de um ponto fixo no cromossomo doador, chamado **origem (O)**, e continua de um modo linear. O ponto O hoje é conhecido como o sítio no qual o plasmídeo F é inserido. Quanto mais longe o gene estiver de O, mais tarde ele será transferido para F⁻. O processo de transferência geralmente pára antes que o mais distante seja transferido, e, como resultado, esses genes são incluídos em menos exconjugantes.

Como podemos explicar a segunda propriedade incomum dos cruzamentos Hfr, de que os exconjugantes F⁻ são raramente convertidos em Hfr ou F⁺? Quando Wollman e Jacob permitiram que os cruzamentos Hfr × F⁻ continuassem por pelo menos 2 h antes da separação, eles observaram que, de fato, alguns dos exconjugantes foram convertidos em Hfr. Em outras palavras, a parte de F que confere a habilidade doadora era transmitida, mas em uma frequência muito baixa. A raridade dos exconjugantes Hfr sugeriu que o F inserido foi transmitido como o **último** elemento do cromossomo linear. Podemos resumir a ordem de transmissão com o seguinte mapa genético, no qual a seta indica o sentido da transferência, começando com O:



Assim, quase nenhuma das receptoras F⁻ era convertida, pois o fator fertilidade é o último elemento transmitido e, geralmente, o processo de transmissão terá parado antes de ir tão longe.

O cromossomo Hfr, originalmente circular, desenrola uma cópia de si mesmo que é transferida para a célula F⁻ de modo linear, com o fator F entrando por último.

Rastreamento do tempo de entrada do marcador gera um mapa cromossômico

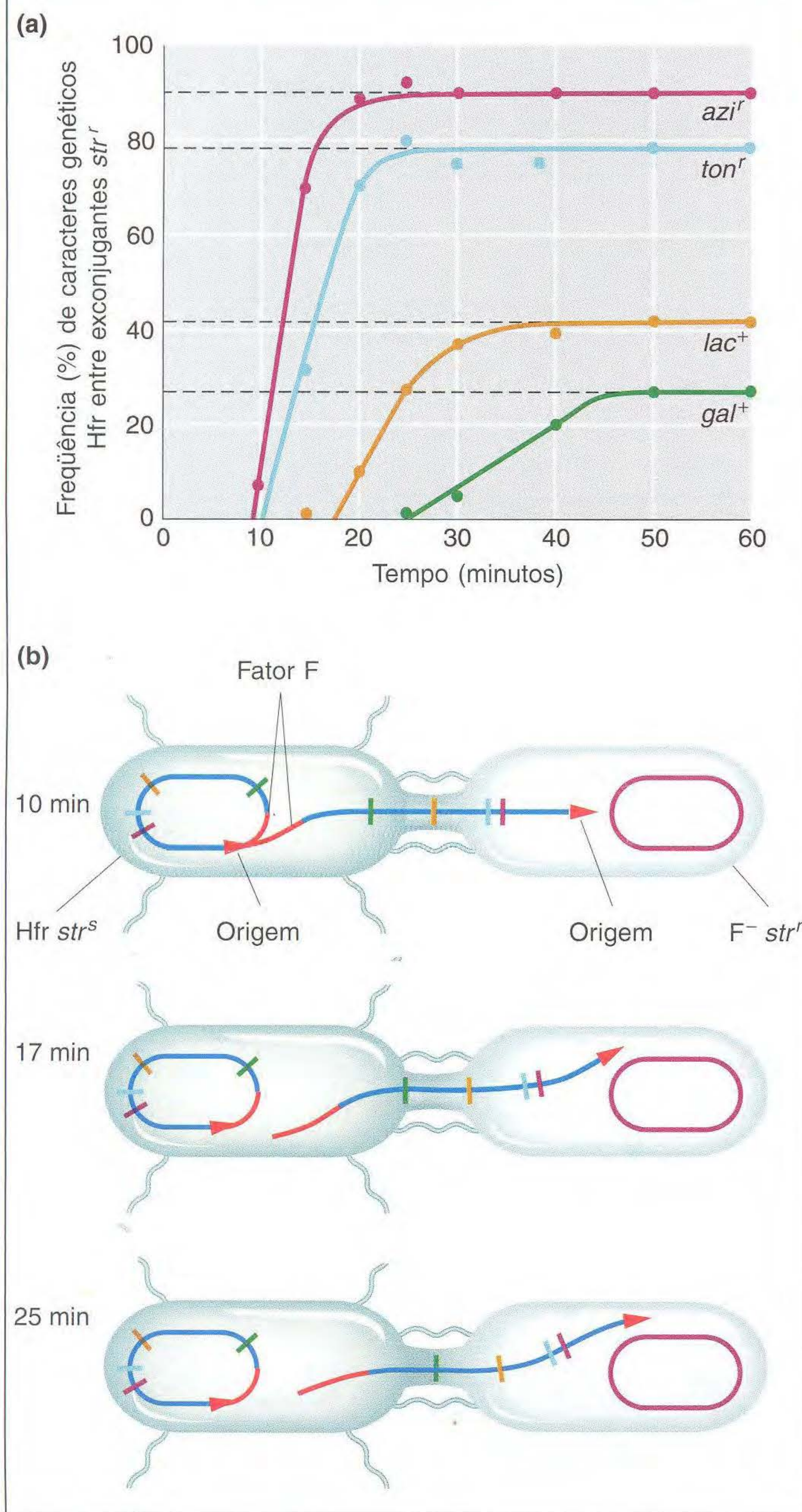


FIG. 5.12 Nesse experimento de conjugação com reprodução interrompida, células F⁻ resistentes a estreptomicina com mutações em *azi*, *ton*, *lac* e *gal* são incubadas por tempos variados com células Hfr que são sensíveis a estreptomicina e levam alelos tipo selvagem para esses genes. (a) Uma plotagem da frequência dos alelos doadores em exconjugantes como função de tempo após a reprodução. (b) Uma visão esquemática da transferência de marcadores (mostrados em cores diferentes) com o passar do tempo. [(a) De E.L. Wollman, F. Jacob, and W. Hayes, *Cold Spring Harbor Symp. Quant. Biol.* 21, 1956, 141.]

Dedução dos sítios de integração de F e circularidade do cromossomo Wollman e Jacob esclareceram como e onde o plasmídeo F se integra para formar uma Hfr, e, fazendo isso, deduziram que o cromossomo é circular. Eles fizeram experi-

mentos de reprodução interrompida com linhagens Hfr diferentes, obtidas separadamente. Significativamente, a ordem de transmissão dos alelos diferia de linhagem a linhagem, como nos exemplos seguintes:

Linhagem Hfr	
H	O <i>tre pro lac pur gal his gli tia</i> F
1	O <i>tre tia gli his gal pur lac pro</i> F
2	O <i>pro tre tia gli his gal pur lac</i> F
3	O <i>pur lac pro tre tia gli his gal</i> F
AB 312	O <i>tia tre pro lac pur gal his gli</i> F

Cada linhagem pode ser considerada um mapa mostrando a ordem dos alelos no cromossomo. À primeira vista, parece haver um embaralhamento aleatório de genes; entretanto, quando alguns dos mapas de Hfr são invertidos, a relação das seqüências torna-se clara.

H	F <i>tia gli his gal pur lac pro tre</i> O
(escrito ao contrário)	
1	O <i>tre tia gli his gal pur lac pro</i> F
2	O <i>pro tre tia gli his gal pur lac</i> F
3	O <i>pur lac pro tre tia gli his gal</i> F
AB 312	F <i>gli his gal pur lac pro tre tia</i> O
(escrito ao contrário)	

A relação de seqüências uma com a outra é explicada se cada mapa for o segmento de um círculo. Isso foi a primeira indicação de que os cromossomos bacterianos são circulares. Além disso, Allan Campbell propôs uma hipótese ótima que explicava os diferentes mapas Hfr. Ele sugeriu que, se F é um anel, então a inserção pode ser um simples crossing entre F e o cromossomo bacteriano (Figura 5.13). Sendo esse o caso, qualquer um dos cromossomos Hfr podia ser gerado pela simples inserção de F no anel no local e orientação apropriados (Figura 5.14).

Várias hipóteses, depois apoiadas, seguiram a proposta de Campbell.

1. Uma ponta do fator F integrado seria a **origem**, onde começa a transferência do cromossomo Hfr. O **término** seria na outra ponta de F.
2. A orientação na qual F é inserido determinaria a ordem de entrada dos alelos doadores. Se o círculo contém os genes A, B, C e D, então a inserção entre A e D daria a ordem ABCD ou DCBA, dependendo da orientação. Veja as diferentes orientações das inserções na Figura 5.14.

Como é possível para F se integrar em sítios diferentes? Se o DNA F tiver uma região homóloga a qualquer das várias regiões no cromossomo bacteriano, qualquer uma delas poderia agir como uma região de pareamento na qual o pareamento poderia ser seguido por um crossing. Essas regiões de homologia hoje são conhecidas como principalmente segmentos de elementos de transposição chamados *seqüências de inserção*. Para uma explicação completa das seqüências de inserção, veja Capítulo 14.

Assim, o fator fertilidade existe em dois estados:

1. O estado de plasmídeo: Como um elemento citoplasmático livre, F é facilmente transferido para receptoras F⁻.

Um único crossing insere F em um locus específico, que então determina a ordem da transferência gênica

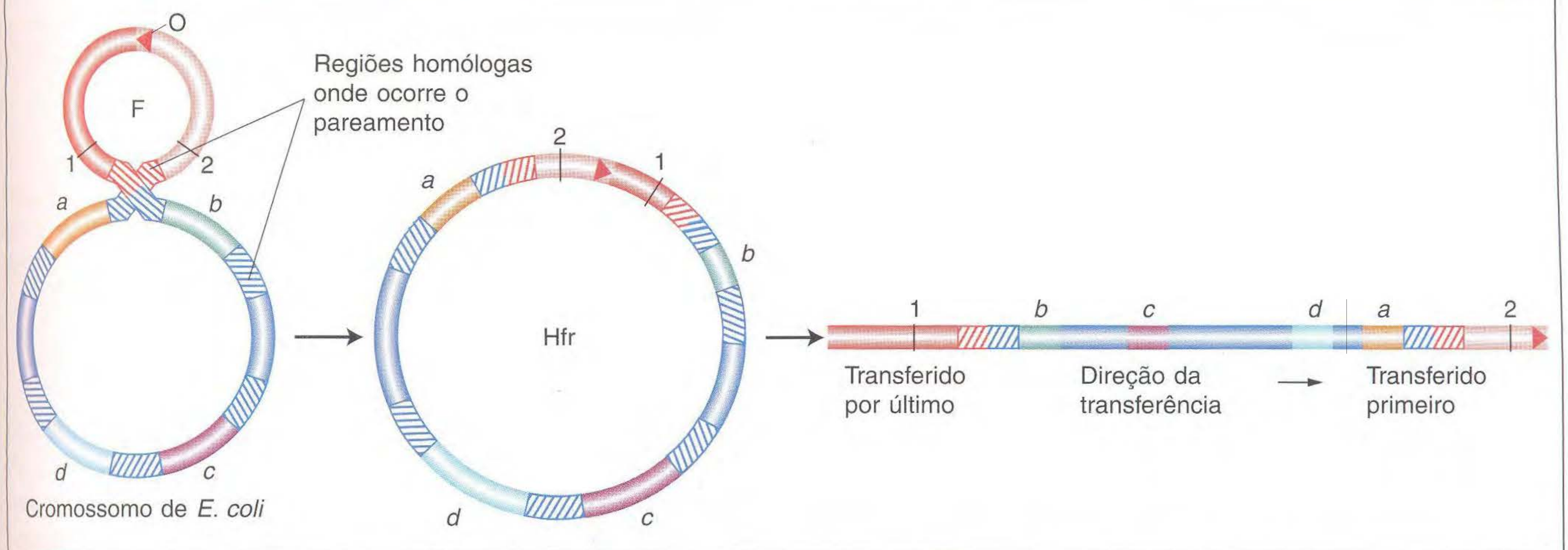


FIG. 5.13 A inserção de F cria uma célula Hfr. Marcadores 1 e 2 hipotéticos são mostrados em F para representar a direção da inserção. A origem (O) é o ponto de mobilização onde ocorre a inserção no cromossomo de *E. coli*; a região de pareamento é homóloga à região no cromossomo de *E. coli*; de a a d representam genes no cromossomo de *E. coli*. As regiões de pareamento (tracejadas) são idênticas no plasmídeo e no cromossomo. Elas são derivadas de elementos móveis chamados *seqüências de inserção* (veja Capítulo 14). Nesse exemplo, a célula Hfr criada pela inserção de F transferiria seus genes na ordem a, d, c, b.

2. O estado integrado: Como uma parte contígua de um cromossomo circular, F é transmitido apenas bem tarde na conjugação.

O ciclo de conjugação de *E. coli* é resumido na Figura 5.15.

Mapeamento de cromossomos bacterianos

Mapeamento cromossômico de ampla escala usando o tempo de entrada Wollman e Jacob perceberam que a construção de mapas de ligação pelos resultados de reprodução interrompida seria fácil, usando como medida de “distância”

O sítio de integração de F determina a ordem de transferência gênica em Hfr

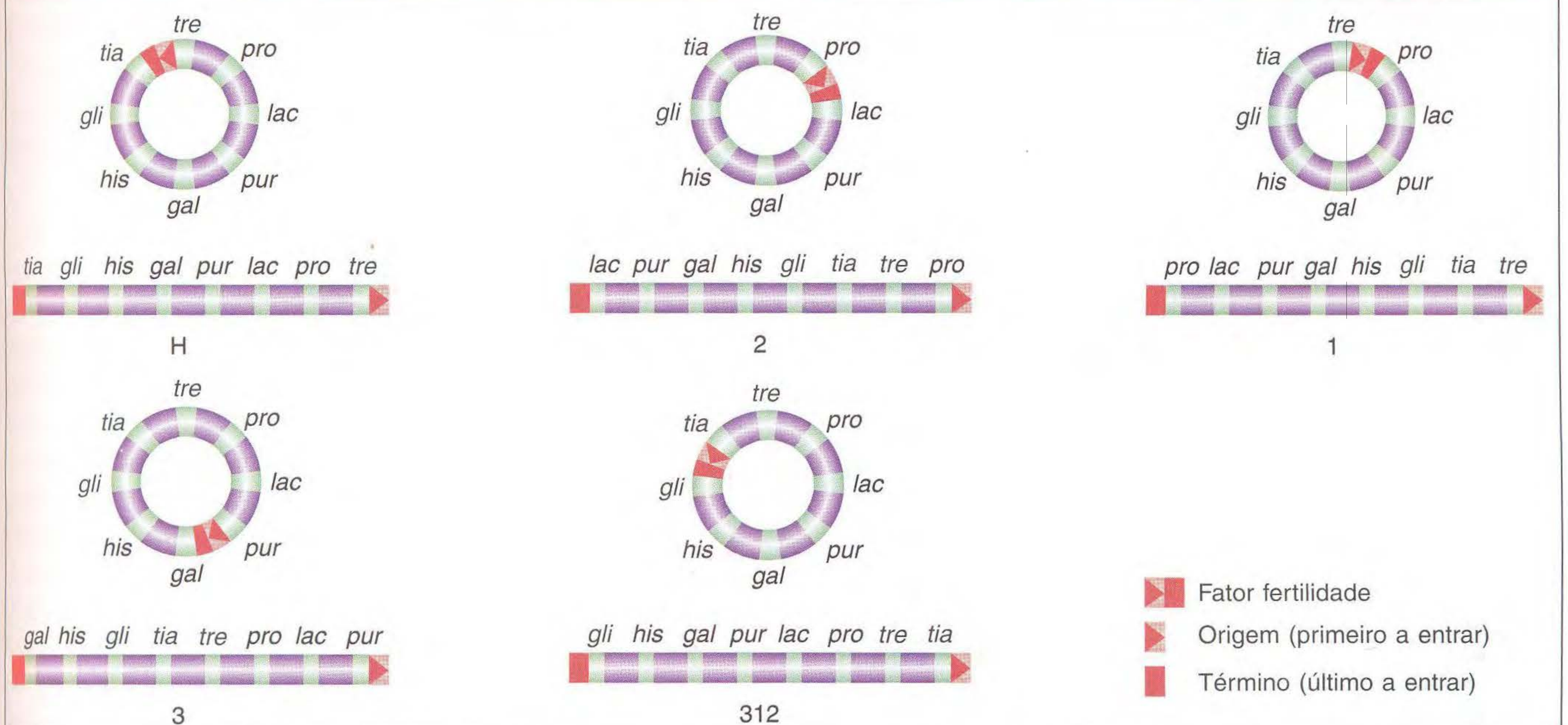


FIG. 5.14 As cinco linhagens Hfr de *E. coli* mostradas têm cada uma pontos de inserção e orientação diferentes no plasmídeo F. Todas as linhagens têm a mesma ordem de genes no cromossomo de *E. coli*. A orientação do fator F determina qual gene entra primeiro na célula receptora. O gene mais próximo do término entra por último.

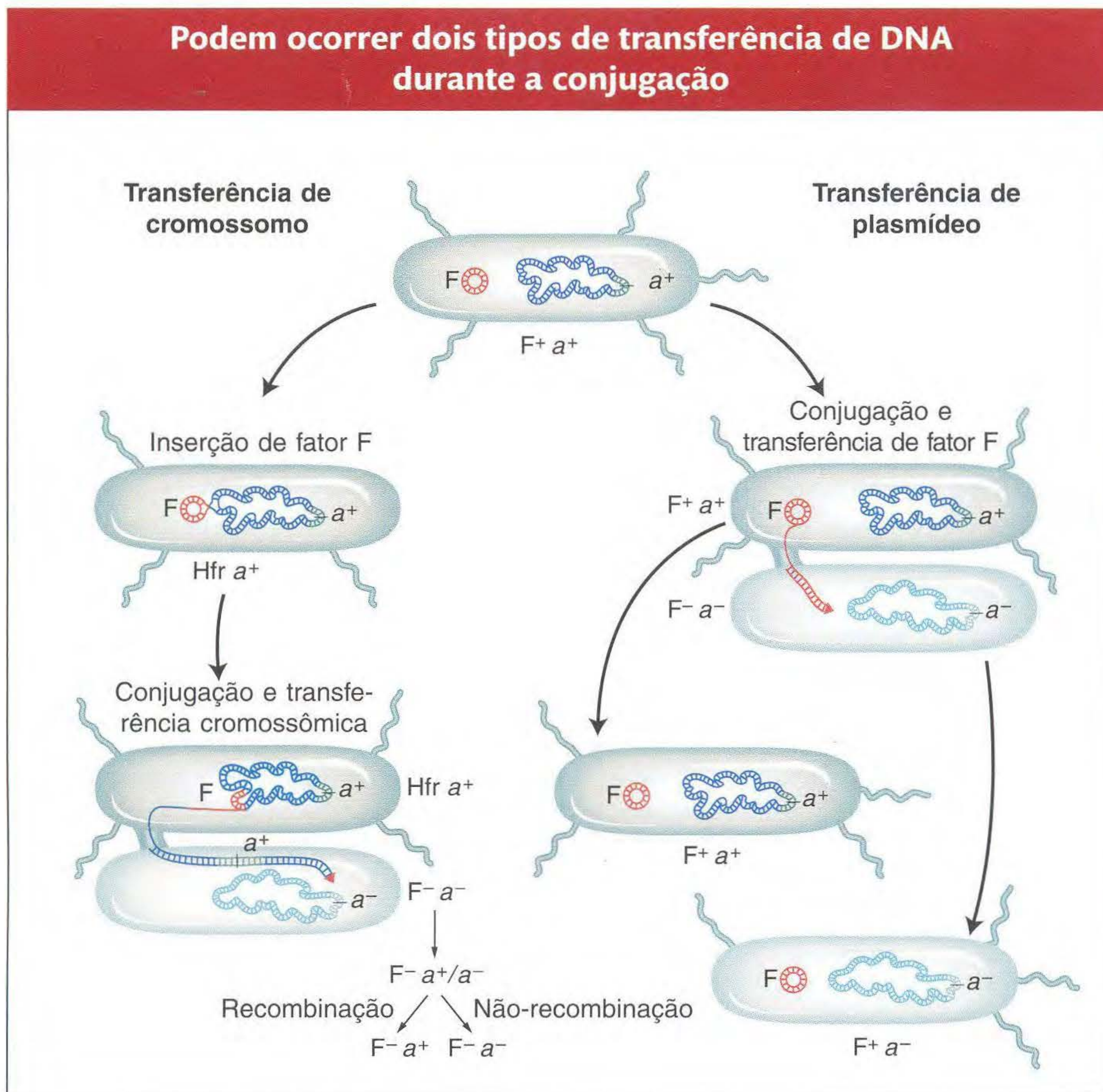


FIG. 5.15 A conjugação pode ocorrer pela transferência parcial de um cromossomo contendo o fator F ou pela transferência de um plasmídeo F que permanece como uma entidade separada.

os tempos nos quais os alelos doadores primeiro apareciam após a reprodução. As unidades de distância de mapa nesse caso são minutos. Assim, se b^+ começa a entrar na célula F^- 10 min após a^+ começar a entrar, então a^+ e b^+ distam 10 unidades. Como os mapas eucarióticos baseados em crossings, esses mapas de ligação foram originalmente puras construções genéticas. Na época em que eles foram originalmente criados, não havia modo de testar sua base física.

Mapeamento cromossômico em escala fina usando frequência de recombinação Para uma exconjugante obter genes doadores como uma característica permanente de seu genoma, o fragmento doador deve se recombinar com o cromossomo receptor. Entretanto, note que o mapeamento por tempo de entrada não é baseado em frequência de recombinantes. De fato, as unidades são minutos, não FR. Entretanto, a frequência de recombinantes pode ser usada para um tipo de escala mais fina de mapeamento em bactérias, um método que veremos agora.

Primeiro, precisamos compreender algumas características especiais do evento de recombinação em bactérias. Lembre-se de que a recombinação não ocorre entre dois genomas inteiros, como ocorre em eucariontes. Em contraste, ocorre entre um genoma *completo*, de F^- , chamado de **endogenoto**, e um *incompleto*, derivado da doadora Hfr e

chamada **exogenoto**. A célula nesse estágio tem duas cópias de um segmento de DNA: uma cópia é parte do endogenoto e a outra cópia é parte do exogenoto. Assim, nesse estágio, a célula é um diplóide *parcial*, chamado de **merozigoto**. A genética bacteriana é genética de merozigoto. Um único crossing em um merozigoto quebraria o anel e, assim, não produziria recombinantes viáveis, como mostrado na Figura 5.16. Para manter o círculo intacto, deve haver um número par de crossings. Um número par de crossings produz um cromossomo circular intacto e um fragmento. Embora tais eventos recombinantes sejam representados de modo resumido como crossings duplos, o real mecanismo molecular é um pouco diferente, mais como uma invasão do endogenoto por uma seção interna do exogenoto. O outro produto do “crossing duplo”, o fragmento, é geralmente perdido no crescimento celular subsequente. Assim, apenas um dos produtos recíprocos da recombinação sobrevive. Portanto, outra característica única da recombinação bacteriana é que temos que esquecer os produtos de troca recíproca na maioria dos casos.

A recombinação durante a conjugação resulta de um evento tipo crossing duplo, que origina recombinantes recíprocos dos quais apenas um sobrevive.

Um único crossing não pode produzir um recombinante viável

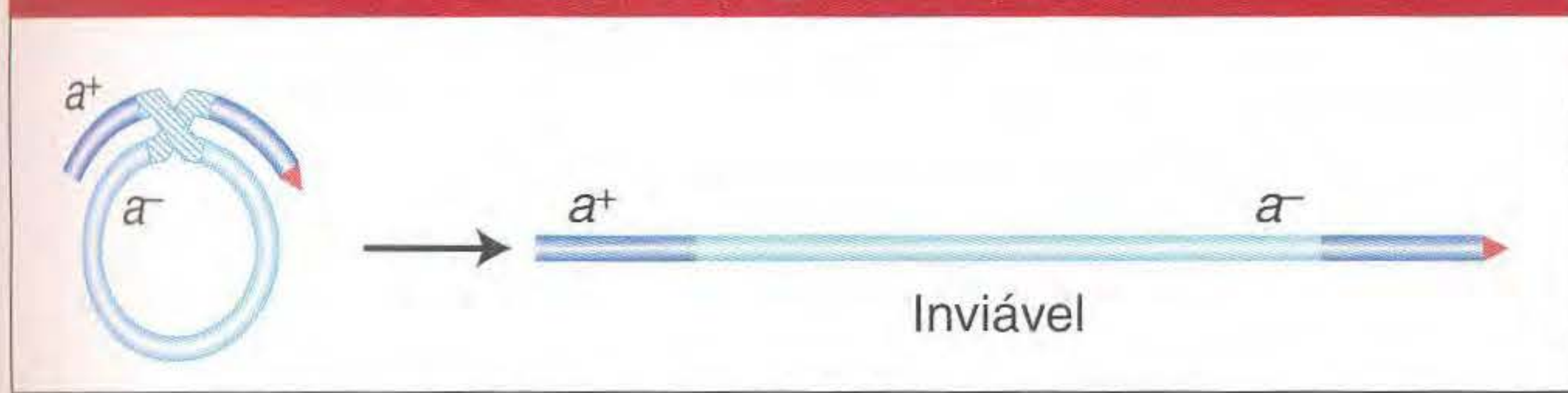
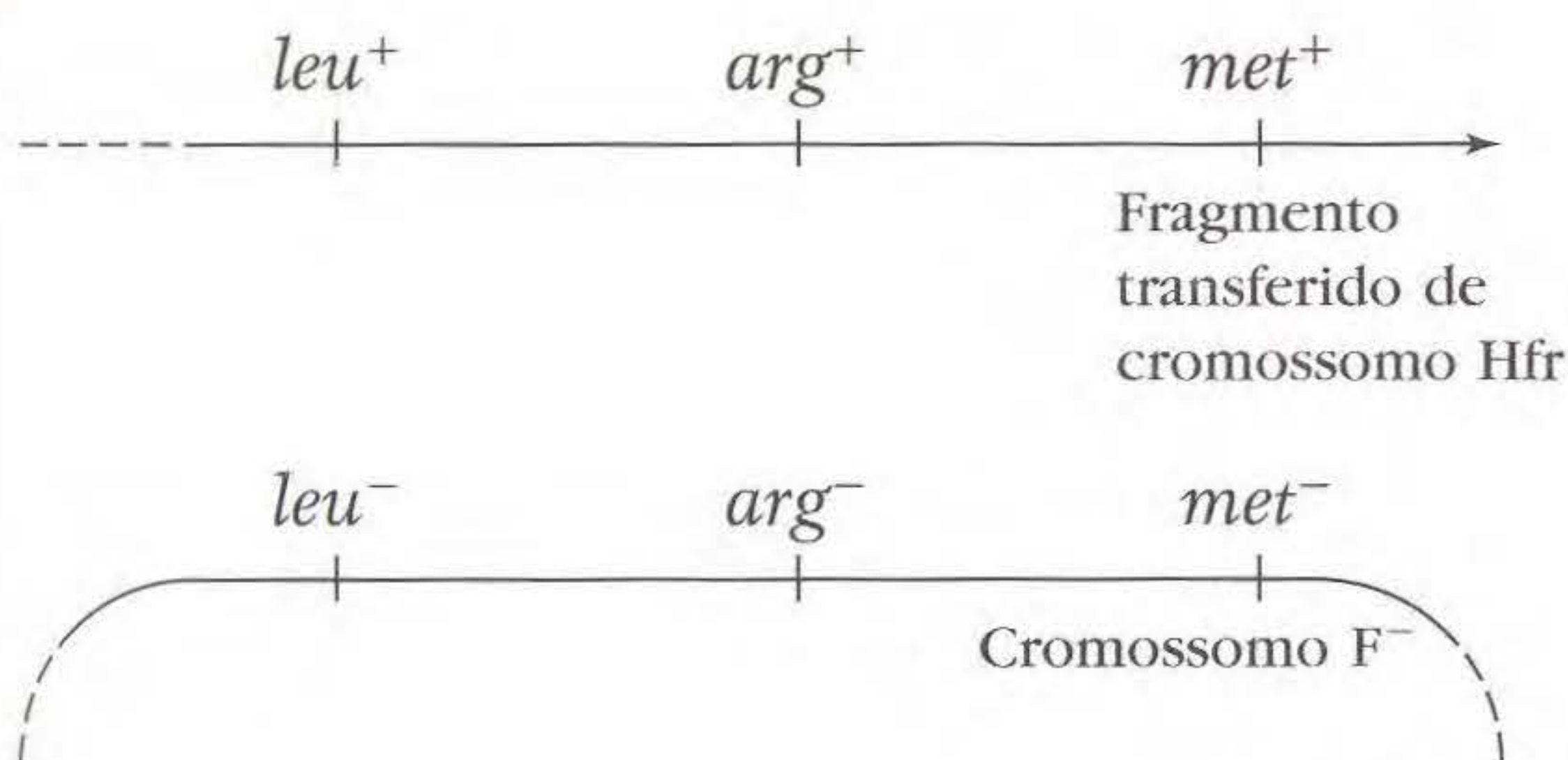


FIG. 5.16 Um único crossing entre o exogenoto e o endogenoto em um merozigoto levaria a um cromossomo linear, parcialmente diplóide, que não sobreviveria.

Com essa compreensão, podemos examinar o mapeamento de recombinação. Suponha que queremos calcular as distâncias de mapa separando três loci próximos: *met*, *arg* e *leu*. Para examinar a recombinação desses genes, precisamos de "tríbridos", exconjugantes que receberam todos os três marcadores doadores. Suponha que um experimento de reprodução interrompida mostrou que a ordem é *met*, *arg*, *leu*, com *met* transferido primeiro e *leu* por último. Para obtermos um tríbrido, precisamos do merozigoto diagramado aqui:



Para obtermos esse merozigoto, devemos primeiro selecionar exconjugantes estáveis portando o *último* alelo doador, que, nesse caso, é leu^+ . Por quê? Porque, se selecionarmos o último marcador, então saberemos que tais células em algum estágio devem ter contido os marcadores anteriores também, isto é, arg^+ e met^+ .

A meta agora é contar as frequências dos crossings em diferentes locais. Note que, agora, temos uma situação diferente da análise de conjugação interrompida. No mapeamento por conjugação interrompida, medimos o tempo de entrada dos loci individuais. Para ser estavelmente herdado, cada marcador tem que se recombinar com o cromossomo receptor por um duplo crossing que o inclua. Entretanto, na análise da frequência de recombinantes, selecionamos especificamente tríbridos como ponto de partida, e agora temos que considerar as várias combinações possíveis dos três alelos doadores que podem ser inseridos por duplo crossing nos vários intervalos. Sabemos que leu^+ deve ter entrado e se inserido porque nós o selecionamos, mas os recombinantes leu^+ que selecionamos podem ou não ter incorporado os outros marcadores doadores, dependendo de onde ocorreu o duplo crossing. Assim, o procedimento é, primeiro, selecionar exconjugantes leu^+ e, então, isolar e testar uma grande amostra deles para ver quais dos outros marcadores foram integrados. Vejamos um exemplo. No cruzamento $Hfr\ met^+ arg^+ leu^+ str^s \times F^- met^- arg^- leu^- str^r$, selecionaríamos recombinantes leu^+ que,

então, examinaríamos quanto aos alelos arg^+ e met^+ , chamados de **marcadores não-selecionados**. A Figura 5.17 mostra os tipos de eventos de crossing duplo esperados. Um crossing deve estar do lado esquerdo do marcador *leu*, e o outro deve estar do lado direito. Suponhamos que os exconjugantes leu^+ são dos seguintes tipos e frequências:

$leu^+ arg^- met^-$	4%
$leu^+ arg^+ met^-$	9%
$leu^+ arg^+ met^+$	87%

Os crossings duplos necessários para produzir esses genótipos são mostrados na Figura 5.17. Os primeiros dois tipos são a chave porque requerem um crossing entre *leu* e *arg*, no primeiro caso, e entre *arg* e *met*, no segundo. Assim, as frequências relativas desses tipos correspondem aos tamanhos dessas duas regiões. Nós concluiríamos que a região *leu-arg* tem 4 unidades de mapa e que *arg-met* tem 9 unidades de mapa.

Em um cruzamento tal como o descrito, um tipo de potenciais recombinantes, com o genótipo $leu^+ arg^- met^+$, requer quatro crossings ao invés de dois (veja a parte inferior da Figura 5.17). Esses recombinantes raramente são recuperados, porque sua frequência é muito baixa comparada com a de outros tipos de recombinantes.

Plasmídeos F que levam fragmentos genômicos

O fator F em linhagens Hfr é geralmente bem estável em sua posição de inserção. Entretanto, ocasionalmente um fator F claramente sai de de um cromossomo por uma reversão do processo de recombinação que o inseriu em primeiro lugar. As duas regiões homólogas de pareamento em ambos os lados fazem novo pareamento, e ocorre um crossing over para liberar o plasmídeo F. Entretanto, às vezes a saída não é limpa, e o plasmídeo leva com ele uma parte do cromossomo bacteriano. Um plasmídeo F levando o DNA genômico bacteriano é chamado de **plasmídeo F' (F primo)**.

A primeira evidência desse processo veio de experimentos, em 1959, por Edward Adelberg e François Jacob. Uma de suas observações importantes foi de uma Hfr cujo fator F foi integrado perto do locus *lac*⁺. Começando com essa linhagem $Hfr\ lac^+$, Jacob e Adelberg encontraram um F^+ derivativo que, nos cruzamentos, transferiu *lac*⁺ para receptores $F^- lac^-$ em uma alta frequência. (Esses transferentes podiam ser detectados plaqueando em meio sem lactose.) O *lac*⁺ transferido não é incorporado ao cromossomo receptor principal, que nós sabemos retém o alelo *lac*⁻ porque esses exconjugantes $F^+ lac^+$ ocasionalmente dão origem a células-filhas $F^- lac^-$, a uma frequência de 1×10^{-3} . Assim, o genótipo desses receptores parece ser $F' lac^+ / F^- lac^-$. Em outras palavras, os exconjugantes *lac*⁺ parecem carregar um plasmídeo F' com um pedaço do cromossomo doador incorporado. A origem desse plasmídeo F' é mostrada na Figura 5.18. Note que a excisão defeituosa ocorre porque há uma outra região homóloga perto que parecia com a original. A F', em nosso exemplo, é chamada F' *lac* porque o pedaço do cromossomo hospedeiro que é captado tem em si o gene *lac*. Os fatores F' foram encontra-

A geração de vários recombinantes por crossing em regiões diferentes

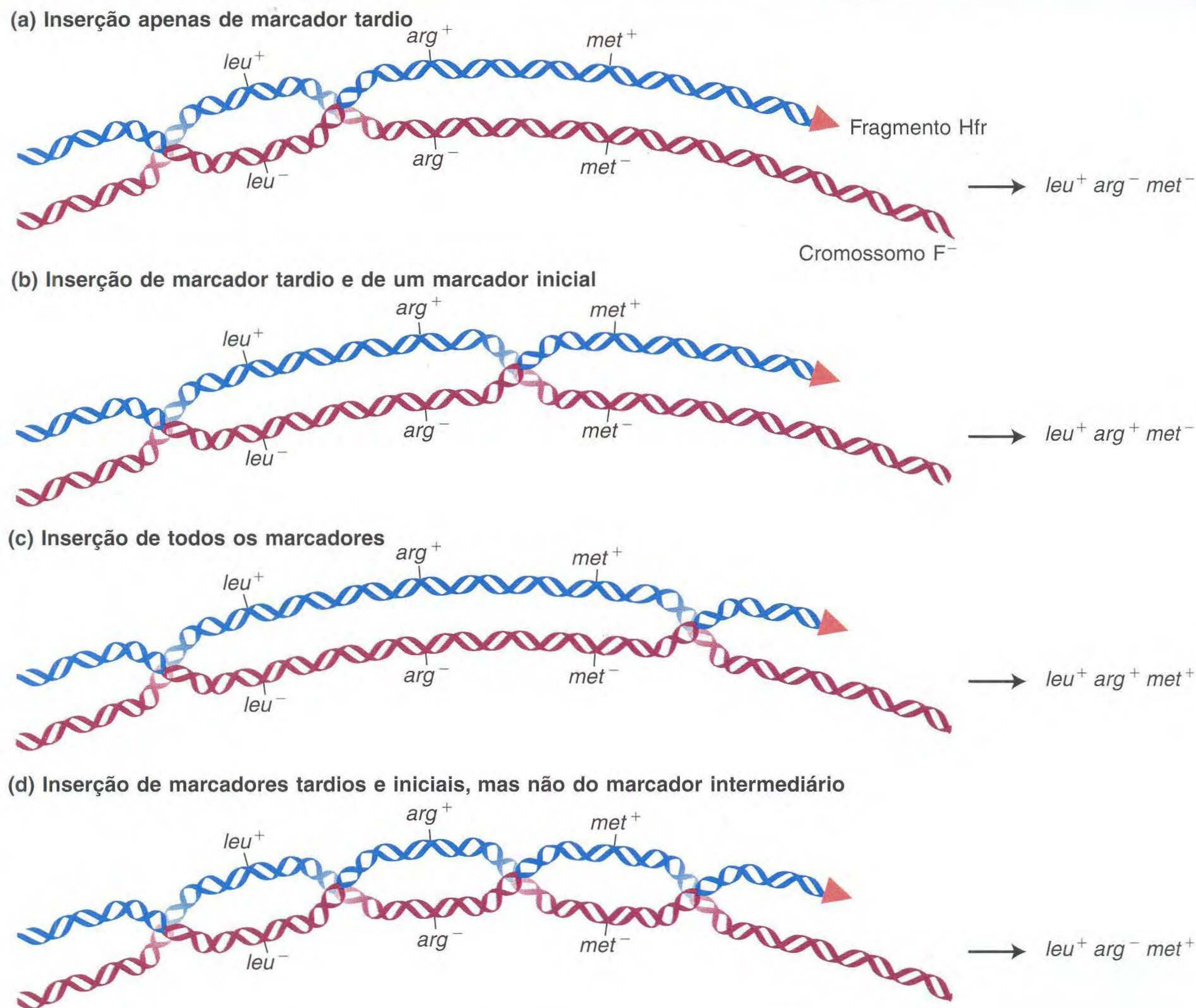


FIG. 5.17 O diagrama mostra como os genes podem ser mapeados por recombinação em *E. coli*. Nos exconjugantes, a seleção é feita para merozigotos portando o marcador *leu*⁺, que é doado posteriormente. Os marcadores iniciais (*arg*⁺ e *met*⁺) podem ou não ser inseridos, dependendo do sítio onde ocorre a recombinação entre o fragmento Hfr e o cromossomo F⁻. As frequências dos eventos diagramados nas partes a e b são usadas para obter os tamanhos relativos das regiões *leu-arg* e *arg-met*. Note que, em cada caso, apenas o DNA inserido no cromossomo F⁻ sobrevive; o outro fragmento é perdido.

dos levando muitos genes cromossômicos diferentes, e foram denominados de acordo. Por exemplo, os fatores F' levando *gal* ou *trp* são chamados de F' *gal* e F' *trp*, respectivamente. Como as células F *lac*⁺/F⁻ *lac*⁻ são de fenótipo Lac⁺, nós sabemos que *lac*⁺ é dominante em relação a *lac*⁻.

Diplóides parciais feitos com o uso de linhagens F' são úteis por alguns aspectos da genética, tais como o estudo da dominância ou de interação de alelo. Algumas linhagens F' podem levar grandes partes (até um quarto) do cromossomo bacteriano.

O DNA de um plasmídeo F' é parte do fator F e parte do genoma bacteriano. Como os plasmídeos F, os plasmídeos F' são transferidos rapidamente. Eles podem ser usados para estabelecer diplóides parciais para estudos de dominância bacteriana e interação de alelos.

Plasmídeos R

Uma alarmante propriedade das bactérias patogênicas primeiro veio à luz por estudos em hospitais japoneses nos anos de 1950. A disenteria bacteriana é causada por bactérias do gênero *Shigella*. Essa bactéria foi inicialmente sensível a uma ampla gama de antibióticos que foram usados para controlar a doença. Nos hospitais japoneses, entretanto, *Shigella* isolada de pacientes com disenteria mostrou-se simultaneamente resistente a muitas dessas drogas, incluindo penicilina, tetraciclina, sulfanilamida, estreptomicina e cloranfenicol. Essa resistência a múltiplas drogas foi herdada como um único pacote genético, e pode ser transmitida de modo infeccioso, não só para outras linhagens sensíveis de *Shigella*, mas também para outras espécies correlatas de bactérias. Esse talento, que se assemelha à mobilidade do plasmídeo F de *E. coli*, é extraordinariamente útil para bactérias patogênicas porque

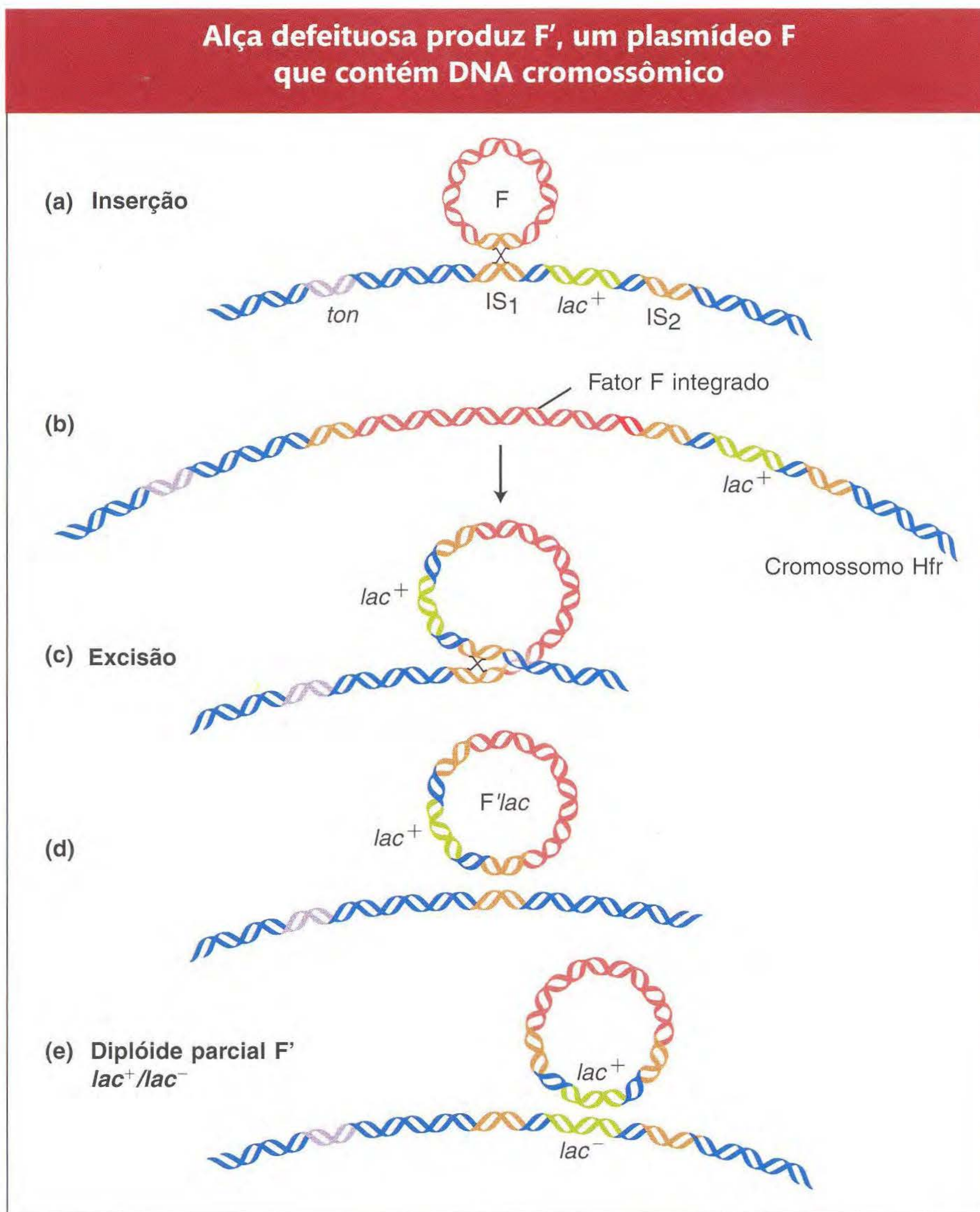


FIG. 5.18 Um fator F pode captar o DNA cromossômico quando ele sai de um cromossomo. (a) F é inserido em uma linhagem Hfr em um elemento repetitivo identificado como IS_1 (seqüência de inserção 1) entre os alelos ton e lac^+ . (b) O fator F inserido. (c) "Alça" anormal por crossing com um elemento diferente, IS_2 , para incluir o locus lac . (d) A partícula lac^+ F' resultante. (e) O diplóide parcial $F' lac^+ / F^- lac^-$ produzido pela transferência da partícula $F' lac^+$ para uma receptora $F^- lac^-$. [De G.S. Stent and R. Calendar, *Molecular Genetics*, 2nd ed. Copyright 1978 by W.H. Freeman and Company.]

a resistência pode se expandir rapidamente por uma população. Entretanto, suas implicações para a ciência médica são terríveis porque a doença bacteriana subitamente se torna resistente ao tratamento por uma grande gama de drogas.

Do ponto de vista do geneticista, entretanto, o mecanismo mostrou-se interessante e é útil na engenharia genética. Os vetores que levam essas múltiplas resistências mostraram ser um outro grupo de plasmídeos chamados **plasmídeos R**. Eles são rapidamente transferidos na conjugação celular, como o plasmídeo F em *E. coli*.

De fato, os plasmídeos R em *Shigella* foram os primeiros de muitos elementos genéticos similares a serem descobertos. Todos existem no estado de plasmídeo no citoplasma. Esses elementos foram encontrados levando muitos tipos de genes em bactérias. O Quadro 5.2 mostra algumas das características que podem surgir de plasmídeos. A Figura 5.19 mostra um exemplo de um plasmídeo isolado da indústria de laticínios.

Os derivados modificados dos plasmídeos R, tais como pBR 322 e pUC (veja Capítulo 20), tornaram-se os vetores preferidos para clonagem molecular do DNA de todos os

Quadro 5.2 Determinantes Genéticos Vindos de Plasmídeos

Características	Exemplos de plasmídeos
Fertilidade	F, R1, Col
Produção de bacteriocina	Col E1
Resistência a metal pesado	R6
Produção de enterotoxina	Ent
Metabolismo de cânfora	Cam
Tumorigenicidade em plantas	T1 (em <i>Agrobacterium tumefaciens</i>)

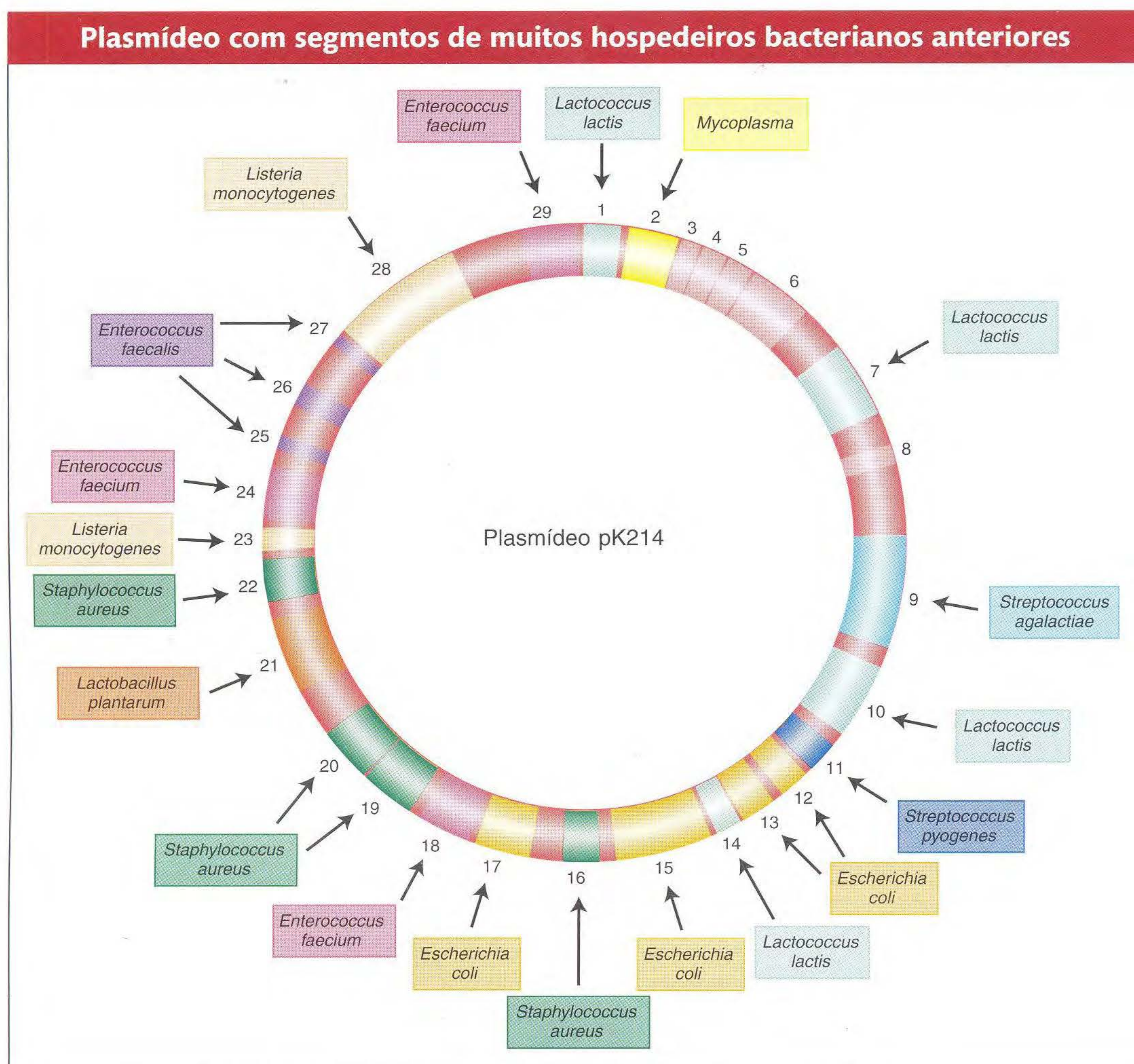


FIG. 5.19 O diagrama mostra as origens dos genes do plasmídeo pK214 de *Lactococcus lactis*. Os genes são de muitas bactérias diferentes. [Dados do Quadro 1 em V. Perreten, F. Schwart, L. Cresta, M. Boeglin, G. Dasen, and M. Teuber, *Nature* 389, 1997, 801-802.]

organismos. Os genes em um plasmídeo R que conferem resistência podem ser usados como marcadores para seguir o movimento dos vetores entre as células.

Nos plasmídeos R, os alelos para resistência a antibióticos em geral estão contidos dentro de uma unidade chamada *transposon* (Figura 5.20). Os transposons são segmentos únicos de DNA que podem mover-se para sítios diferentes no genoma, um processo chamado de transposição. (Os mecanismos para transposição, que ocorrem na maioria das espécies estudadas, serão detalhados no Capítulo 14.) Quando um transposon, no genoma, se move para um novo local, ele pode ocasionalmente levar vários tipos de genes, incluindo alelos para resistência a drogas, e levá-los para seus novos locais como passageiros. Às vezes, um transposon leva um alelo para resistência a drogas para um plasmídeo, criando um plasmídeo R. Como os plasmídeos F, muitos plasmídeos R são conjugativos; em outras palavras, eles são efetivamente transmitidos para uma célula receptora durante a conjugação. Mesmo os plasmídeos R que são não-conjugativos e nunca deixam suas próprias células podem doar seus alelos R para um plasmídeo conjugativo por transposição. Assim, através dos plasmídeos, os alelos para resistência a antibióticos

podem se difundir rapidamente por uma população de bactérias. Embora a dispersão de plasmídeos R seja uma estratégia

Um plasmídeo R com genes de resistência levados em um transposon

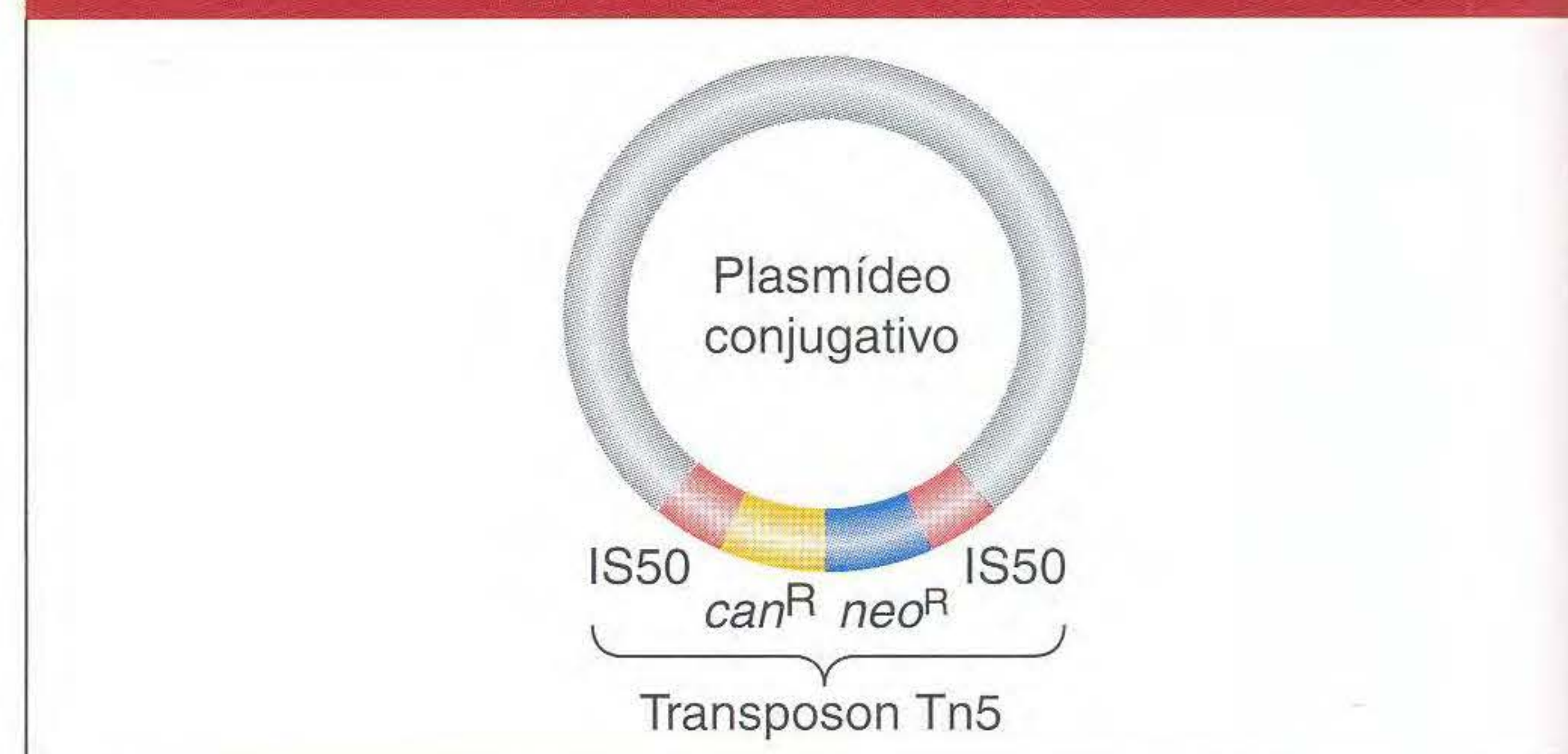


FIG. 5.20 Um transposon tal como Tn5 pode adquirir vários genes resistentes a drogas (nesse caso, aqueles para resistência a drogas canamicina e neomicina) e os transmite rapidamente para um plasmídeo, levando à transferência infecciosa de genes de resistência como um pacote. A sequência de inserção 50 (IS50) forma flancadores de TN5.

efetiva para a sobrevivência das bactérias, ela representa um problema importante para a prática médica, como mencionado antes, pois as populações de bactérias rapidamente se tornam resistentes a qualquer nova droga antibiótica que é inventada e aplicada aos humanos.

► 5.3 Transformação Bacteriana

Algumas bactérias podem captar fragmentos de DNA do meio externo, e tal captação constitui outro modo pelo qual as bactérias podem trocar seus genes. A fonte do DNA pode ser outras células da mesma espécie ou células de outras espécies. Em alguns casos, o DNA foi liberado de células mortas; em outros casos, o DNA foi secretado de células bacterianas vivas. O DNA captado integra-se ao cromossomo receptor. Se esse DNA é de um genótipo diferente do da receptora, o genótipo da receptora pode tornar-se permanentemente alterado, um processo adequadamente chamado de **transformação**.

A transformação foi descoberta na bactéria *Streptococcus pneumoniae* em 1928, por Frederick Griffith. Mais tarde, em 1944, Oswald T. Avery, Colin M. MacLeod e Maclyn McCarty demonstraram que o “princípio transformante” era o DNA. Ambos os resultados são marcos na elucidação da natureza molecular dos genes. Consideramos esse trabalho em mais detalhes no Capítulo 7.

O DNA transformante é incorporado ao cromossomo bacteriano por um processo análogo aos eventos de dupla recombinação observados nos cruzamentos $Hfr \times F^-$. Note, entretanto, que, na *conjugação*, o DNA é transferido de

uma célula viva para outra por meio de um contato íntimo, enquanto, na *transformação*, pedaços isolados do DNA externo são captados por uma célula através da parede celular e membrana plasmática. A Figura 5.21 mostra um modo pelo qual ocorre esse processo.

A transformação tem sido uma ferramenta útil em várias áreas da pesquisa bacteriana porque o genótipo de uma linhagem pode ser deliberadamente alterado de um modo muito específico transformando-se com um fragmento apropriado de DNA. Por exemplo, a transformação é amplamente usada na engenharia genética. Mais recentemente, descobriu-se que mesmo as células eucarióticas podem ser transformadas, usando-se procedimentos muito similares, e essa técnica é valiosa para modificar células eucarióticas.

Mapeamento cromossômico usando transformação

A transformação pode ser usada para medir o quanto dois genes ligados estão próximos em um cromossomo bacteriano. Quando o DNA (o cromossomo bacteriano) é extraído para experimentos de transformação, é inevitável alguma quebra em pedaços menores. Se dois genes doadores estão situados muito próximos no cromossomo, há uma boa chance de, às vezes, eles serem levados no mesmo pedaço do DNA transformante. Assim, ambos serão captados, causando uma **transformação dupla**. Contrariamente, se os genes estão bem separados no cromossomo, eles mais provavelmente serão levados em segmentos transformantes separados. Um genoma poderia possivelmente captar ambos os segmentos independentemente, criando um transformante duplo, mas esse resultado não é provável. Assim, em genes bem separados, a frequência de transformantes duplos será igual ao produto das frequências de transformantes únicos. Portanto, deve ser possível testar a proximidade de ligação testando a saída pela regra do produto. Em outras palavras, se os genes são ligados, então a proporção de transformantes duplos será maior que o produto das frequências de transformantes únicos.

Infelizmente, a situação é mais complicada por vários fatores, sendo o mais importante que nem todas as células em uma população de bactérias são competentes para serem transformadas. Entretanto, ao final deste capítulo, você poderá aprimorar suas habilidades na análise da transformação em um dos problemas, que supõe que 100% das células receptoras são competentes.

As bactérias podem captar fragmentos de DNA do meio ambiente. Dentro da célula, esses fragmentos podem se integrar ao cromossomo.

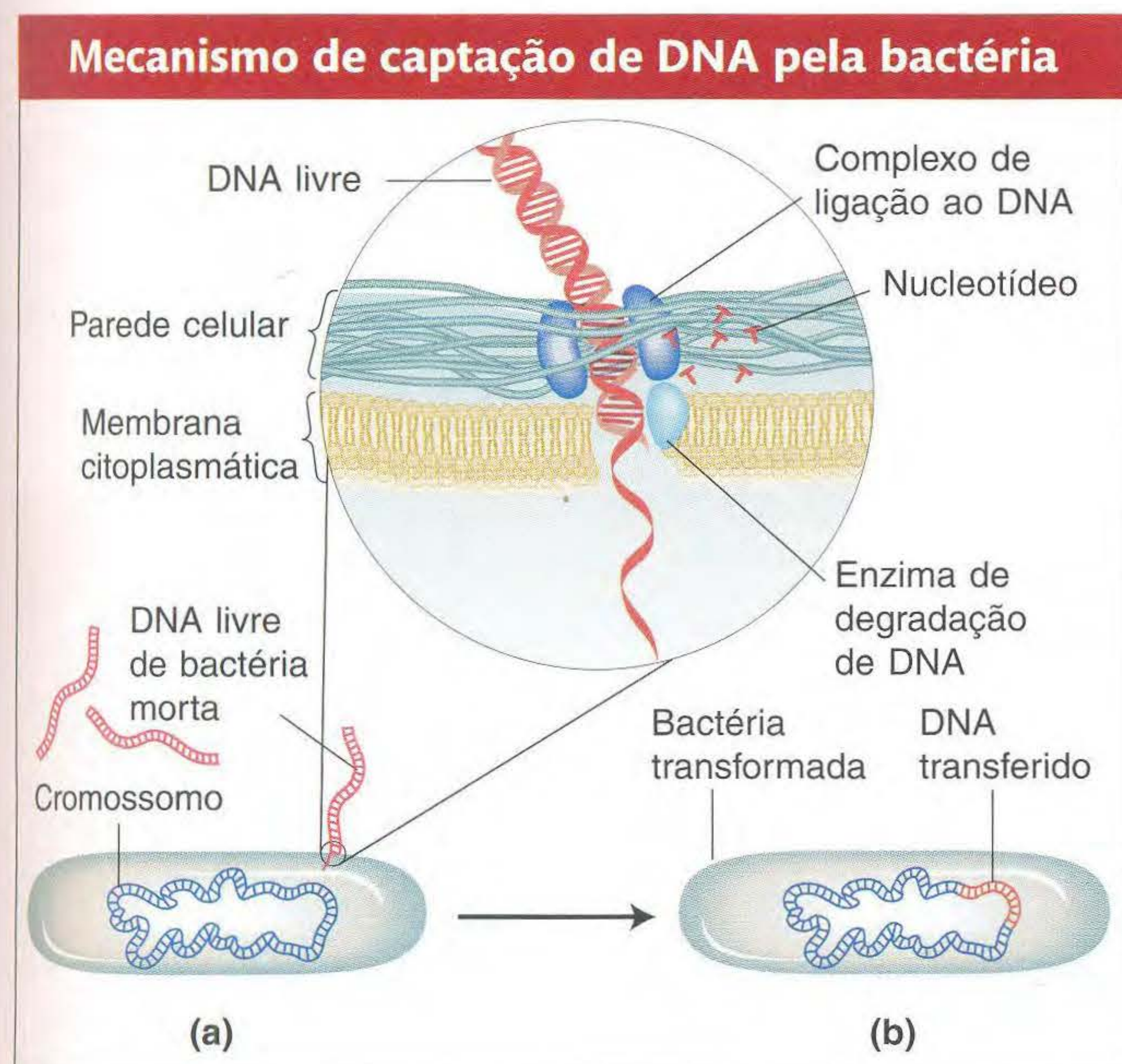


FIG. 5.21 Uma bactéria sofrendo transformação (a) capta DNA livre liberado de uma bactéria morta. À medida que os complexos de ligação ao DNA na superfície da bactéria captam o DNA (*inserto*), as enzimas degradam um filamento em nucleotídeos; um derivado do outro filamento pode integrar-se ao cromossomo bacteriano (b). [De R.V. Miller, "Bacterial Gene Swapping in Nature." Copyright 1998 de Scientific American, Inc. Todos os direitos reservados.]

► 5.4 Genética de Bacteriófagos

A palavra *bacteriófago*, que é um nome para vírus bacterianos, significa “comedor de bactérias”. Esses vírus parasitam e matam a bactéria. O trabalho pioneiro em genética de bacte-

Estrutura e função do fago T4

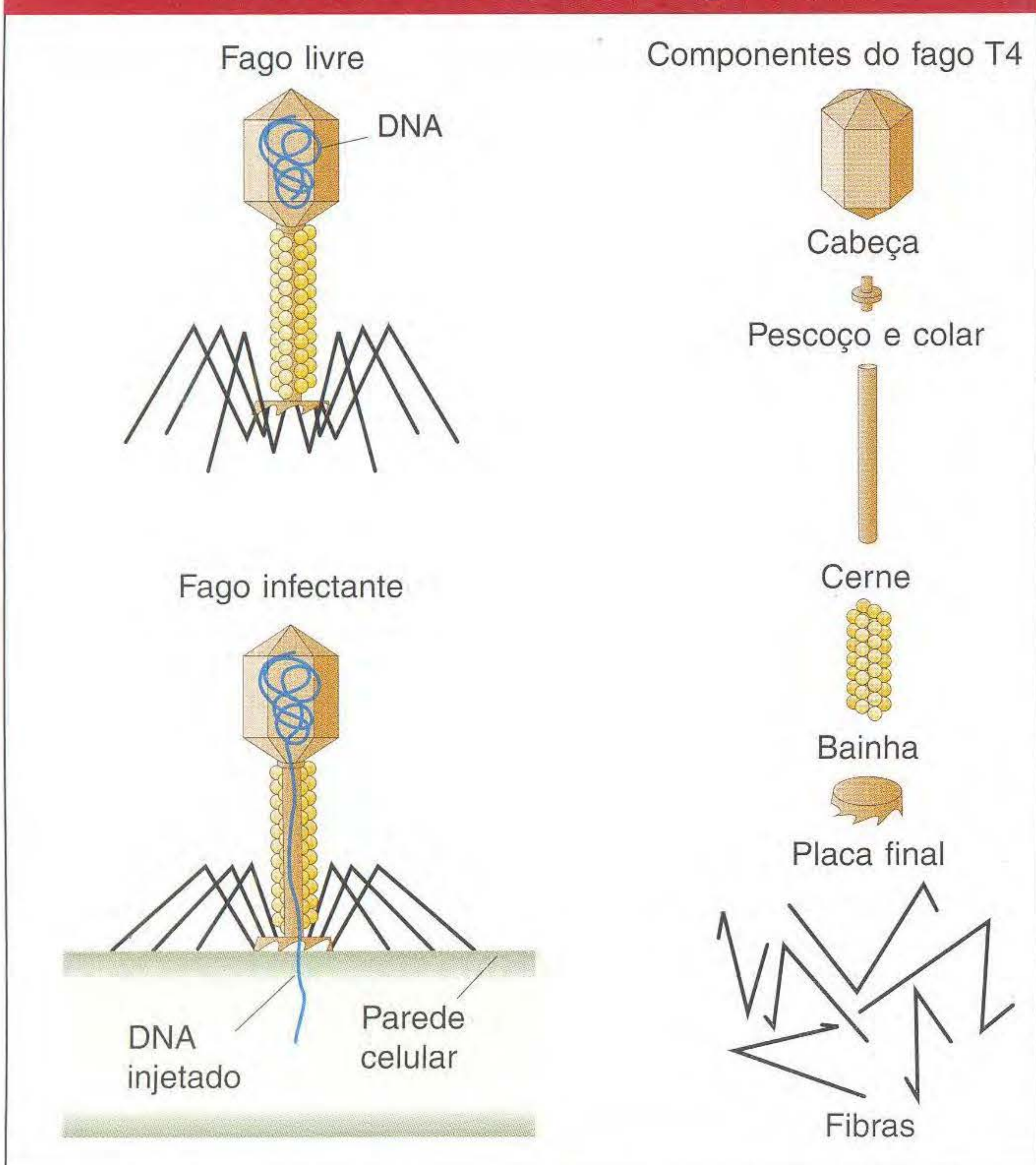


FIG. 5.22 Um fago infectante injeta DNA por sua estrutura cerne na célula. (Esquerda) O bacteriófago T4 é mostrado como um fago livre e, então, no processo de infecção de uma *E. coli*. (Direita) Os principais componentes estruturais de T4. [De R.S. Edgar and R.H. Epstein, "The Genetics of a Bacterial Virus." Copyright 1965 de Scientific American, Inc. Todos os direitos reservados.]

riófagos, na metade do século XX, formou a fundação da pesquisa mais recente sobre vírus causadores de tumor e outros tipos de vírus animais e de plantas. Desse modo, os vírus bacterianos forneceram um importante sistema-modelo.

Esses vírus podem ser usados em dois tipos diferentes de análise genética. Primeiro, dois genótipos distintos de fagos

podem ser cruzados para medir a recombinação e, portanto, mapear o genoma viral. O mapeamento do genoma viral por esse método é o tópico desta seção. Segundo, os bacteriófagos podem ser usados como um modo de juntar genes bacterianos para estudos de ligação e outros estudos genéticos. Estudaremos o uso de fagos em estudos bacterianos na Seção 5.5. Além disso, como veremos no Capítulo 20, os fagos são usados na tecnologia do DNA como portadores, ou vetores, de DNA exógeno. Antes que possamos compreender a genética de fagos, devemos examinar primeiro o ciclo de infecção dos fagos.

Infecção de bactérias por fagos

A maioria das bactérias são suscetíveis ao ataque por bacteriófagos. Um fago consiste em um "cromossomo" de ácido nucleico (DNA ou RNA) circundado por uma capa de moléculas de proteína. Os tipos de fagos são identificados não por nomes de espécies, mas por símbolos — por exemplo, fago T4, fago λ , e assim por diante. As Figuras 5.22 e 5.23 mostram a estrutura do fago T4. Durante a infecção, um fago liga-se a uma bactéria e injeta seu material genético no citoplasma bacteriano, como diagramado na Figura 5.22. Uma micrografia eletrônica do processo é mostrada na Figura 5.24. A informação genética do fago então comanda a maquinaria da célula bacteriana desligando a síntese dos componentes bacterianos e redirecionando a maquinaria de síntese de componentes da bactéria para fazer componentes do fago. As cabeças de fagos recém-formadas são individualmente recheadas com réplicas do cromossomo do fago. Finalmente, são feitos muitos descendentes do fago e são liberados quando a parede celular da bactéria é rompida. Esse processo de abertura é chamado de **lise**. A população de prole de fagos é chamada de **lisado** de fagos.

Como podemos estudar a herança em fagos quando eles são tão pequenos que só são visíveis ao microscópio eletrônico? Nesse caso, não podemos produzir uma colônia visível plaqueando, mas podemos produzir uma manifestação visível de um fago tirando proveito de várias características do fago.

Micrografia eletrônica do fago T4



FIG. 5.23 A ampliação do fago T4 de *E. coli* revela detalhes da cabeça, cauda e fibras da cauda. [Foto de Jack D. Griffith.]

Micrografia eletrônica de infecção de fago



FIG. 5.24 São mostrados bacteriófagos em vários estágios do processo infeccioso, que inclui a ligação e injeção do DNA. [Do Dr. L. Caro/Science Photo Library, Photo Researchers.]

Ciclo de um fago que lisa as células hospedeiras

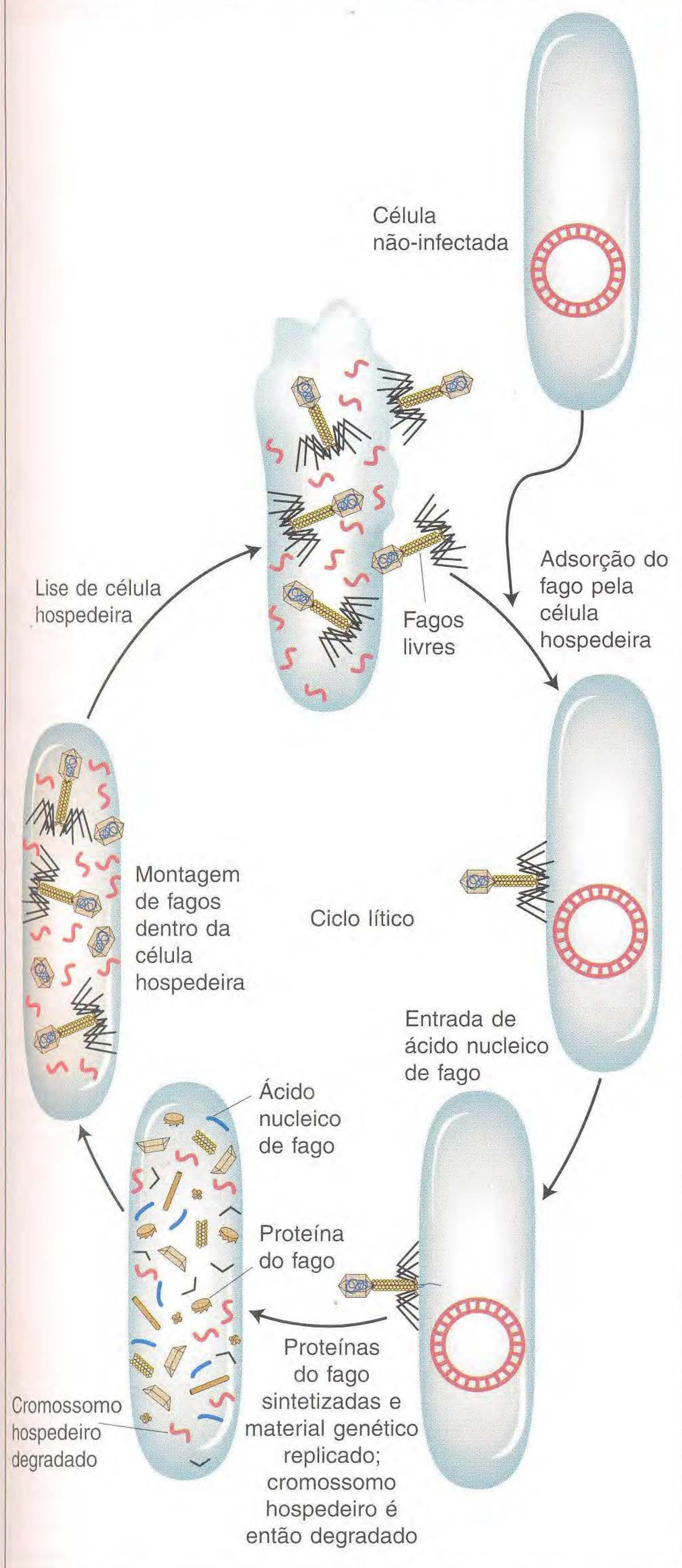


FIG. 5.25 A infecção por um único fago redireciona a maquinaria da célula para fazer fagos de prole, que são liberados na lise. [De J. Darnell, H. Lodish, and D. Baltimore, *Molecular Cell Biology*. Copyright 1986 de W.H. Freeman and Company.]

Vejamos as consequências de um fago infectando uma única bactéria. A Figura 5.25 mostra a sequência de eventos no ciclo infeccioso que leva à liberação da prole de fagos pela célula lisada.

Uma placa é uma área clara na qual todas as bactérias foram lisadas por fagos

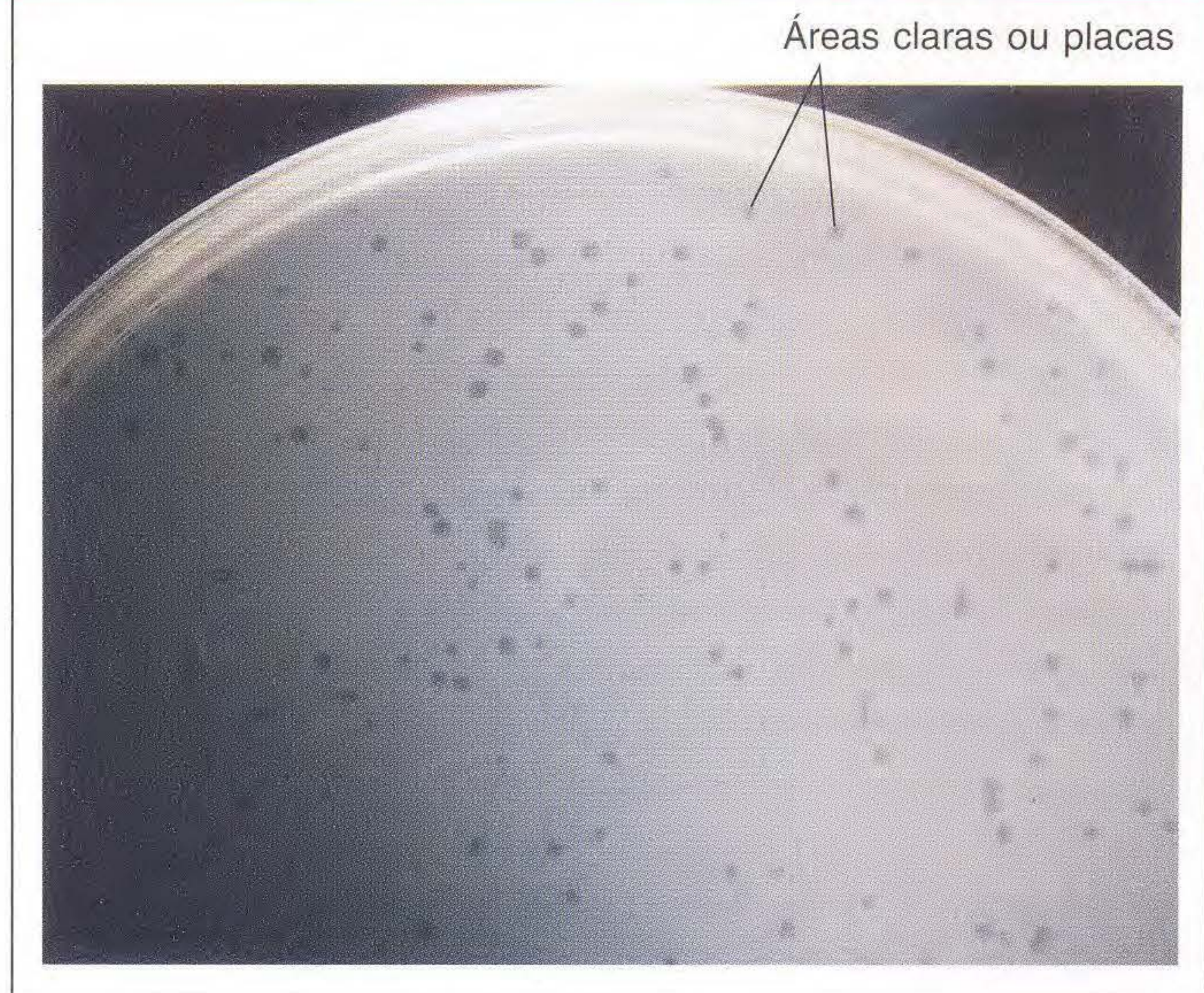


FIG. 5.26 Por infecção repetida e produção de fagos da prole, um único fago produz uma área clara, ou placa, na camada opaca das bactérias. [Barbara Morris, Novagen.]

Após a lise, os fagos da prole infectam as bactérias vizinhas. Esse ciclo é repetido por rodadas progressivas de infecção, e, à medida que os ciclos se repetem, o número de células lisadas aumenta exponencialmente. Dentro de 15 h após uma única partícula de fago infectar uma bactéria, os efeitos são visíveis a olho nu como uma área clara, ou **placa**, na camada opaca de bactérias que cobrem a superfície de uma placa de meio sólido (Figura 5.26). Tais placas podem ser grandes ou pequenas, encrespadas ou pontudas, assim por diante, dependendo do genótipo do fago. Assim, a *morfologia da placa* é uma característica do fago que pode ser analisada a nível genético. Outro genótipo do fago que pode ser analisado geneticamente é a *gama de hospedeiros*, pois os fagos podem diferir no espectro de linhagens bacterianas que eles podem infectar e lisar. Por exemplo, uma linhagem específica de bactérias pode ser imune ao fago 1 mas suscetível ao fago 2.

Mapeamento de cromossomos de fagos usando cruzamentos de fagos

Dois genótipos de fago podem ser cruzados do mesmo modo que cruzamos organismos. Um cruzamento de fago pode ser ilustrado por um cruzamento dos fagos T2 originalmente estudados por Alfred Hershey. Os genótipos das duas linhagens parentais no cruzamento de Hershey eram $h^- r^+ \times h^+ r^-$. Os alelos correspondem aos seguintes fenótipos:

- h^- : pode infectar duas linhagens diferentes de *E. coli* (que podemos chamar de linhagens 1 e 2)
- h^+ : pode infectar apenas a linhagem 1
- r^- : lisa rapidamente as células, produzindo assim grandes placas
- r^+ : lisa lentamente as células, produzindo pequenas placas

Para fazer o cruzamento, a linhagem 1 de *E. coli* é infectada com ambos os genótipos parentais do fago T2. Esse tipo de infecção é chamado de uma **infecção mista** ou uma **infecção dupla** (Figura 5.27). Após um período apropriado de incubação, o lisado de fagos (os fagos da prole) é analisado espalhando-o em uma camada de bactérias composta de uma mistura de linhagens 1 e 2 de *E. coli*. São distinguíveis quatro tipos de placas (Figura 5.28). Grandes placas indicam lise rápida (r^-), e pequenas placas indicam lise lenta (r^+). As placas de fagos com o alelo h^- infectarão ambos os hospedeiros, formando uma placa clara, enquanto placas de fago com o alelo h^+ infectarão apenas um hospedeiro, formando uma placa turva. Assim, os quatro genótipos podem ser facilmente classificados como parentais ($h^- r^+$ e $h^+ r^-$) e recombinantes ($h^+ r^+$ e $h^- r^-$), e uma frequência de recombinantes pode ser calculada do seguinte modo:

$$FR = \frac{(h^+ r^+) + (h^- r^-)}{\text{total de placas}}$$

Se for suposto que os cromossomos recombinantes de fago são lineares, então crossings únicos geram produtos viáveis recíprocos. Entretanto, cruzamentos de fago são sujeitos a algumas complicações analíticas. Primeiro, várias rodadas de troca podem ocorrer dentro do hospedeiro: um recombinante produzido logo após a infecção pode sofrer mais recombinação na mesma célula ou em outros ciclos de infecção. Segundo, a recombinação pode ocorrer entre fagos geneticamente similares, bem como entre tipos diferentes. Assim, se P_1 e P_2 forem genótipos parentais gerais, os cruzamentos de $P_1 \times P_1$ e $P_2 \times P_2$ ocorrem em adição a $P_1 \times P_2$. Por ambos esses motivos, os recombinantes de cruzamentos de fagos são uma consequência de uma *população* de eventos, em vez de eventos de troca definidos de uma etapa. Entretanto, *as outras coisas sendo todas iguais*, o cálculo da FR representa um indicador válido para mapear a distância nos fagos.

Como números imensos de fagos podem ser usados na análise de recombinação de fagos, raros eventos de crossing podem ser detectados. Nos anos de 1950, Seymour Benzer usou raros eventos de crossing para mapear sítios mutan-

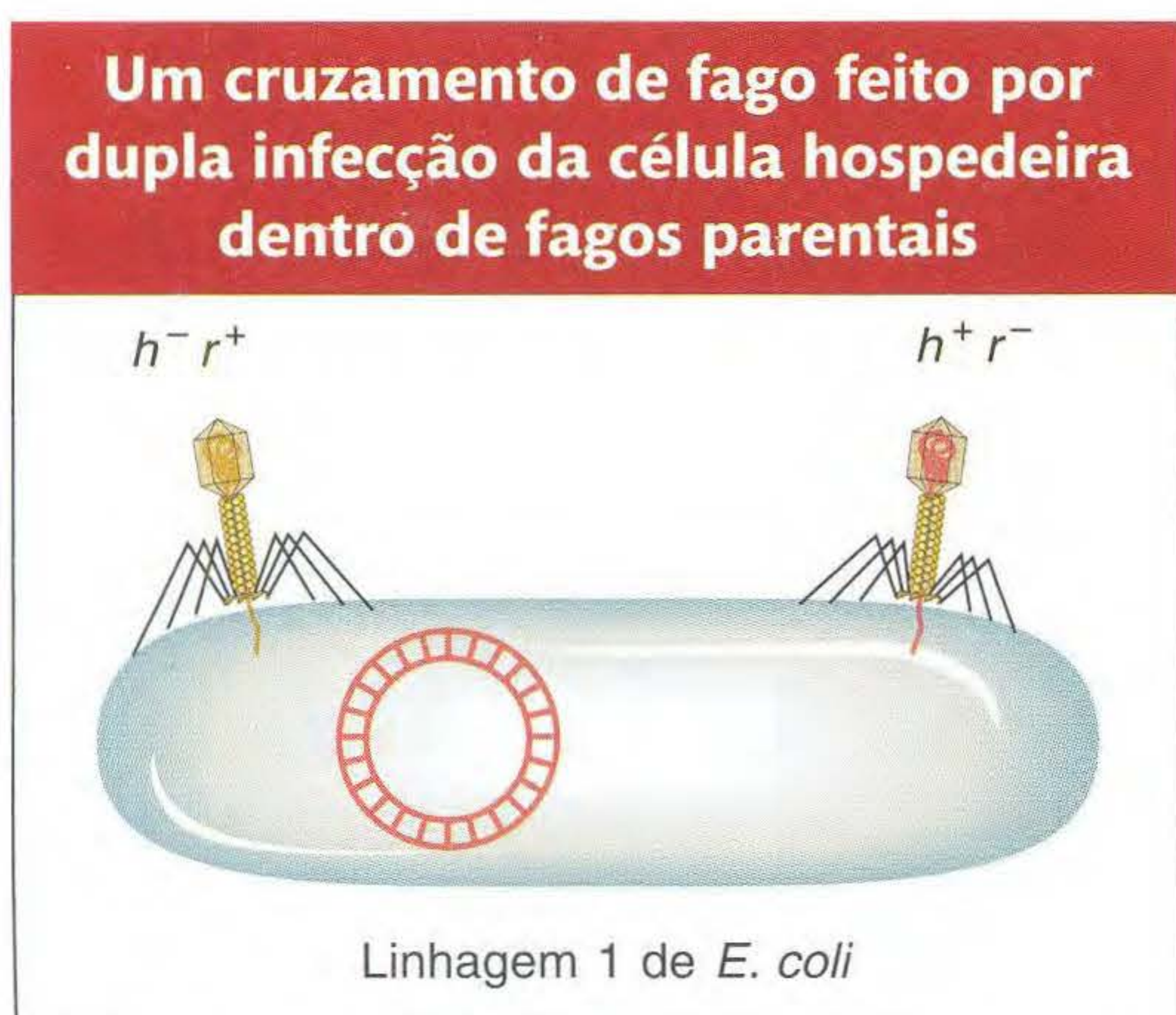


FIG. 5.27

Placas de prole de fagos recombinantes e parentais

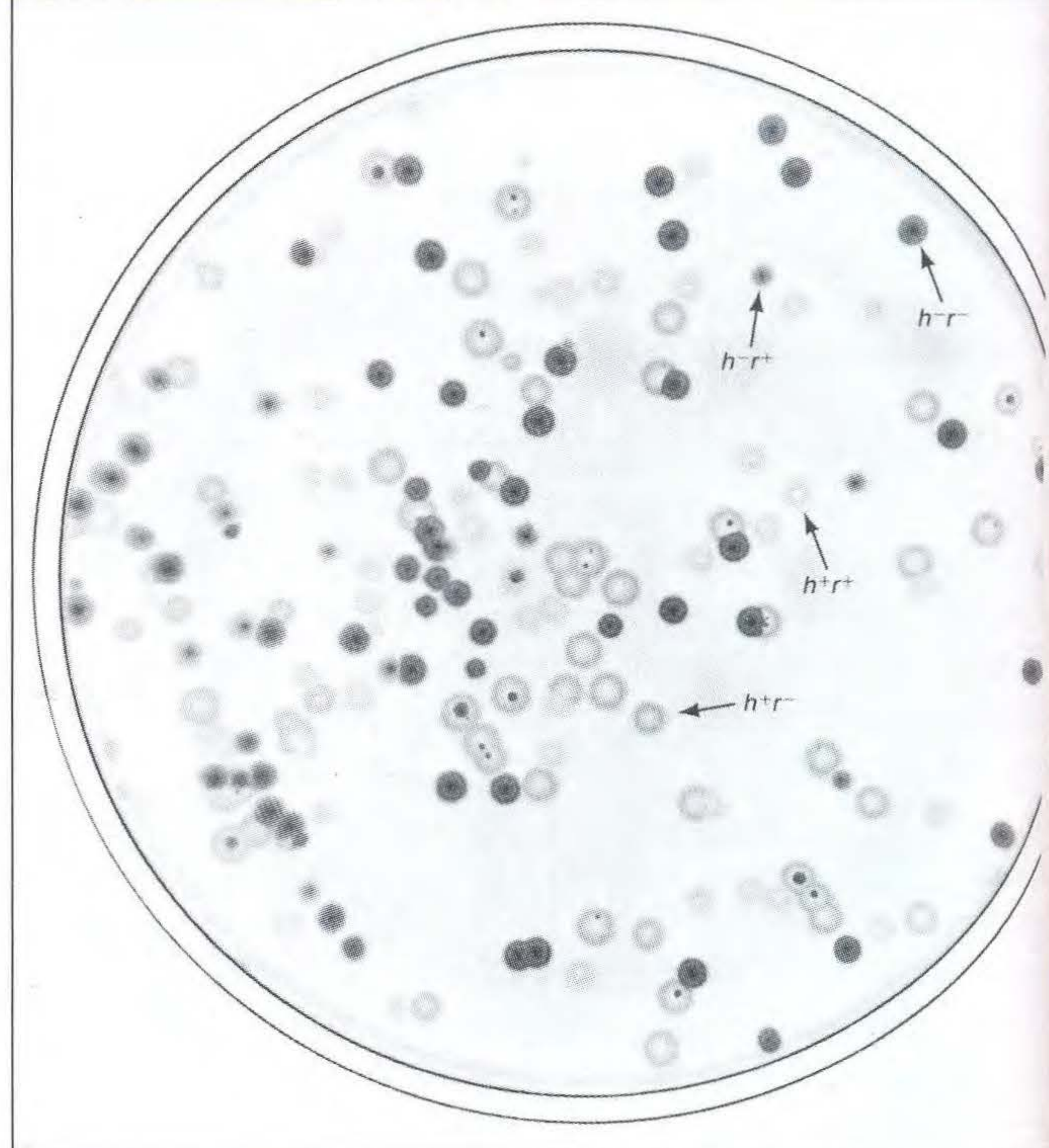
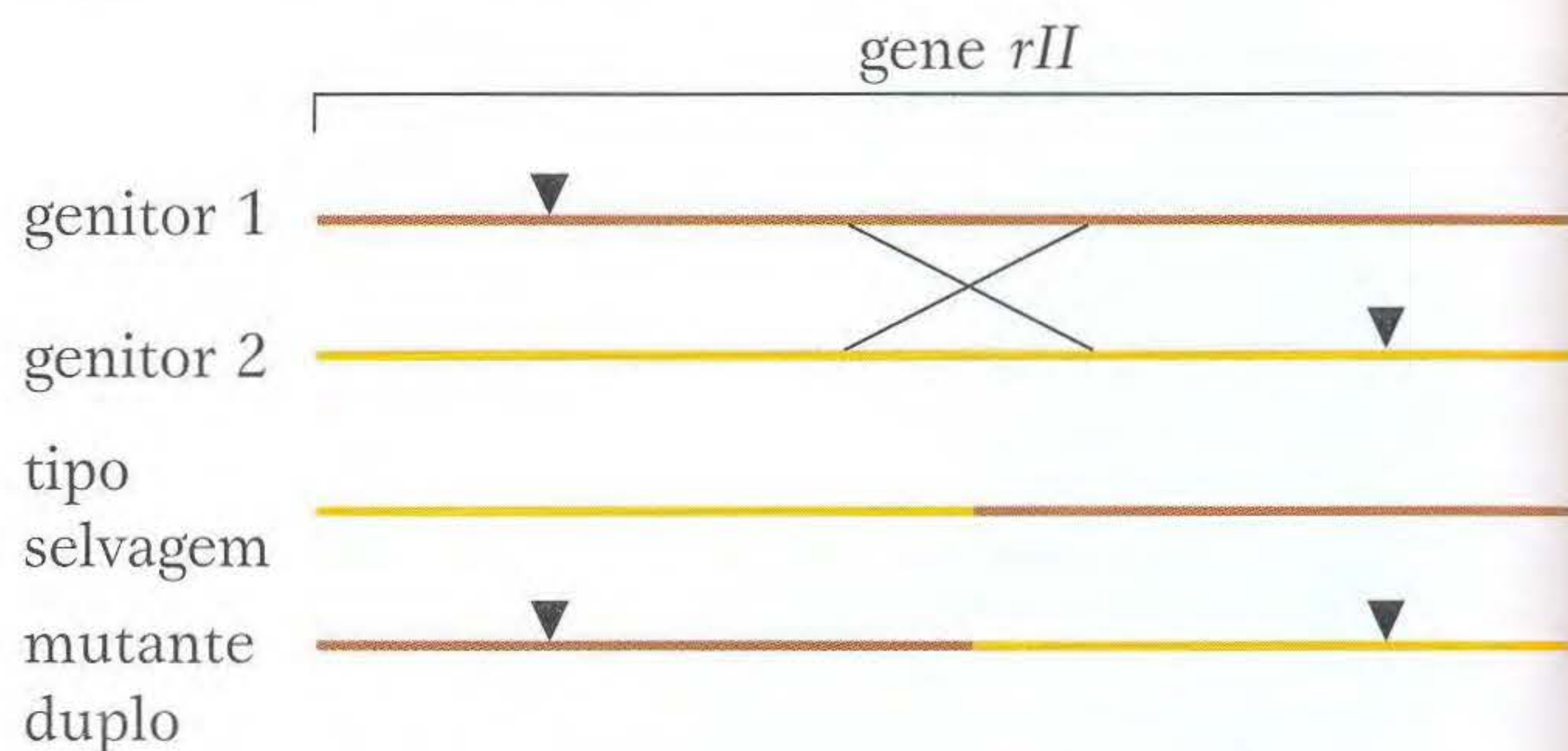


FIG. 5.28 Esses fenótipos de placa foram produzidos pela prole de um cruzamento $h^- r^+ \times h^+ r^-$. Podem ser diferenciados quatro fenótipos de placa, representando dois tipos parentais e dois recombinantes. [De G.S. Stent, *Molecular Biology of Bacterial Viruses*. Copyright 1963 by W.H. Freeman and Company.]

tes *dentro* do gene *rII* do fago T4, um gene que controla a lise. Para diferentes alelos mutantes *rII* que surgem espontaneamente, o sítio mutante está geralmente em posições diferentes dentro do gene. Portanto, quando dois mutantes diferentes *rII* são cruzados, podem ocorrer alguns crossings raros entre os sítios mutantes, produzindo recombinantes tipo selvagem, como mostrado aqui:



À medida que a distância entre dois sítios mutantes aumenta, tal evento de crossing é mais provável. Assim, a frequência de recombinantes rII^+ é uma medida dessa distância dentro do gene. (O produto recíproco é um mutante duplo e indistinguível dos genitores.)

Benzer usou um enfoque inteligente para detectar recombinantes rII^+ muito raros. Ele usou o fato de que os mutan-

tes rII não infectariam uma linhagem de *E. coli* chamada K. Portanto, ele fez o cruzamento rII × rII em outra linhagem, e então plaqueou o lisado de fagos em uma camada da linhagem K. Apenas os recombinantes rII⁺ formarão placas nessa camada. Esse modo de descobrir um raro evento genético (nesse caso, um recombinante) é um **sistema seletivo**: apenas o raro evento desejado pode produzir certo resultado visível. Em contraste, uma **triagem** é um sistema no qual grandes números de indivíduos são visualmente escaneados para procurar a rara “agulha do palheiro”.

O mesmo enfoque pode ser usado para mapear sítios mutantes dentro dos genes de qualquer organismo do qual foi obtido um grande número de células, e para o qual podem ser distintos fenótipos tipo selvagem e mutantes. Entretanto, esse tipo de mapeamento intragênico foi superado pelo advento de métodos químicos baratos para o seqüenciamento do DNA, que identificam diretamente as posições de sítios mutantes.

A recombinação de cromossomos de fagos pode ser estudada juntando-se os cromossomos parentais em uma célula hospedeira por infecção mista. Os fagos da prole podem ser examinados quanto a genótipos parentais e recombinantes.

► 5.5 Transdução

Alguns fagos são capazes de captar genes bacterianos e levá-los de uma bactéria para outra: um processo conhecido como **transdução**. Assim, a transdução junta-se à bateria de modos de transferir material genômico entre bactérias, juntamente com a transferência de cromossomos Hfr, transferência de plasmídeo F' e transformação.

Descoberta da transdução

Em 1951, Joshua Lederberg e Norton Zinder testaram a recombinação na bactéria *Salmonella typhimurium* usando técnicas que foram bem-sucedidas com *E. coli*. Os pesquisadores usaram duas linhagens diferentes: uma era *fen*⁻ *trp*⁺ *tir*⁻, e a outra era *met*⁻ *his*⁻. Não nos preocuparemos com a natureza desses alelos exceto para notar que são todos auxotróficos. Quando ambas as linhagens foram plaqueadas em meio mínimo, não foram observadas células tipo selvagem. Entretanto, após as duas linhagens serem misturadas, surgiram prototróficos tipo selvagem com uma frequência de cerca de 1 em 10⁵. Até aqui, a situação parece similar à da recombinação em *E. coli*.

Entretanto, nesse caso, os pesquisadores também recuperaram recombinantes de um experimento com tubo em U, no qual a conjugação foi evitada por um filtro separando os dois braços. Eles supuseram que algum agente estava levando genes de uma bactéria para outra. Variando o tamanho dos poros no filtro, eles descobriram que o agente responsável pela transferência gênica era do mesmo tamanho que um fago conhecido de *Salmonella*, chamado P22. Além disso, o agente filtrável e P22 eram idênticos, em sensibilidade, ao anti-soro e, em imunidade, a enzimas hidrolíticas. Assim,

Lederberg e Zinder descobriram um novo tipo de transferência gênica, mediado por um vírus. Eles foram os primeiros a chamar esse processo de *transdução*. Como uma raridade no ciclo lítico, as partículas de vírus às vezes captam genes bacterianos e os transferem quando infectam outro hospedeiro. A transdução foi subsequentemente demonstrada em muitas bactérias.

Para compreendermos o processo de transdução, precisamos distinguir dois tipos de ciclo de fago. Os **fagos virulentos** são os que imediatamente lisam e matam o hospedeiro. Os **fagos temperados** podem permanecer dentro da célula hospedeira por um período sem matá-la. Seu DNA ou se integra ao cromossomo do hospedeiro para se replicar com ele, ou replica-se separadamente no citoplasma, como faz um plasmídeo. Um fago integrado ao genoma bacteriano é chamado de **profago**. Uma bactéria que abriga um fago quiescente é chamada de **lisogênica**. Ocasionalmente, o fago quiescente em uma bactéria lisogênica torna-se ativo, replica-se e causa lise espontânea de sua célula hospedeira. Um fago temperado residente confere resistência a infecção por outros fagos desse tipo.

Existem dois tipos de transdução: generalizada e especializada. Os fagos de transdução *generalizada* podem levar qualquer parte do cromossomo bacteriano, enquanto os fagos de transdução especializada levam apenas algumas partes *específicas*.

Os fagos virulentos não podem tornar-se profagos; eles sempre lisam uma célula imediatamente ao entrar. Os fagos temperados podem existir dentro da célula bacteriana como profagos, permitindo que seus hospedeiros sobrevivam como bactérias lisogênicas; eles também são capazes de, ocasionalmente, lisar a bactéria.

Transdução generalizada

Por que mecanismos um fago faz uma **transdução generalizada**? Em 1965, H. Ikeda e J. Tomizawa esclareceram essa questão em alguns experimentos no fago P1 de *E. coli*. Eles observaram que, quando uma célula doadora é lisada por P1, o cromossomo bacteriano é rompido em pequenos pedaços. Ocasionalmente, as novas partículas de fago formadas erroneamente incorporam um pedaço do DNA bacteriano em uma cabeça de fago em lugar do DNA do fago. Esse evento é a origem do fago transdutor.

Um fago levando o DNA bacteriano pode infectar outra célula. Esse DNA bacteriano pode então ser incorporado ao cromossomo da célula receptora por recombinação (Figura 5.29). Como os genes em qualquer uma das partes cortadas do genoma hospedeiro podem ser transduzidos, esse tipo de transdução é necessariamente do tipo transdução generalizada.

Os fagos P1 e P22 ambos pertencem a um grupo de fagos que apresentam transdução generalizada. O DNA de P22 é inserido no cromossomo hospedeiro, enquanto o DNA de P1 permanece livre, como um grande plasmídeo. Entretanto, ambos transduzem por preenchimento defeituoso da cabeça.

Transdução generalizada por incorporação aleatória de DNA bacteriano em cabeças de fagos

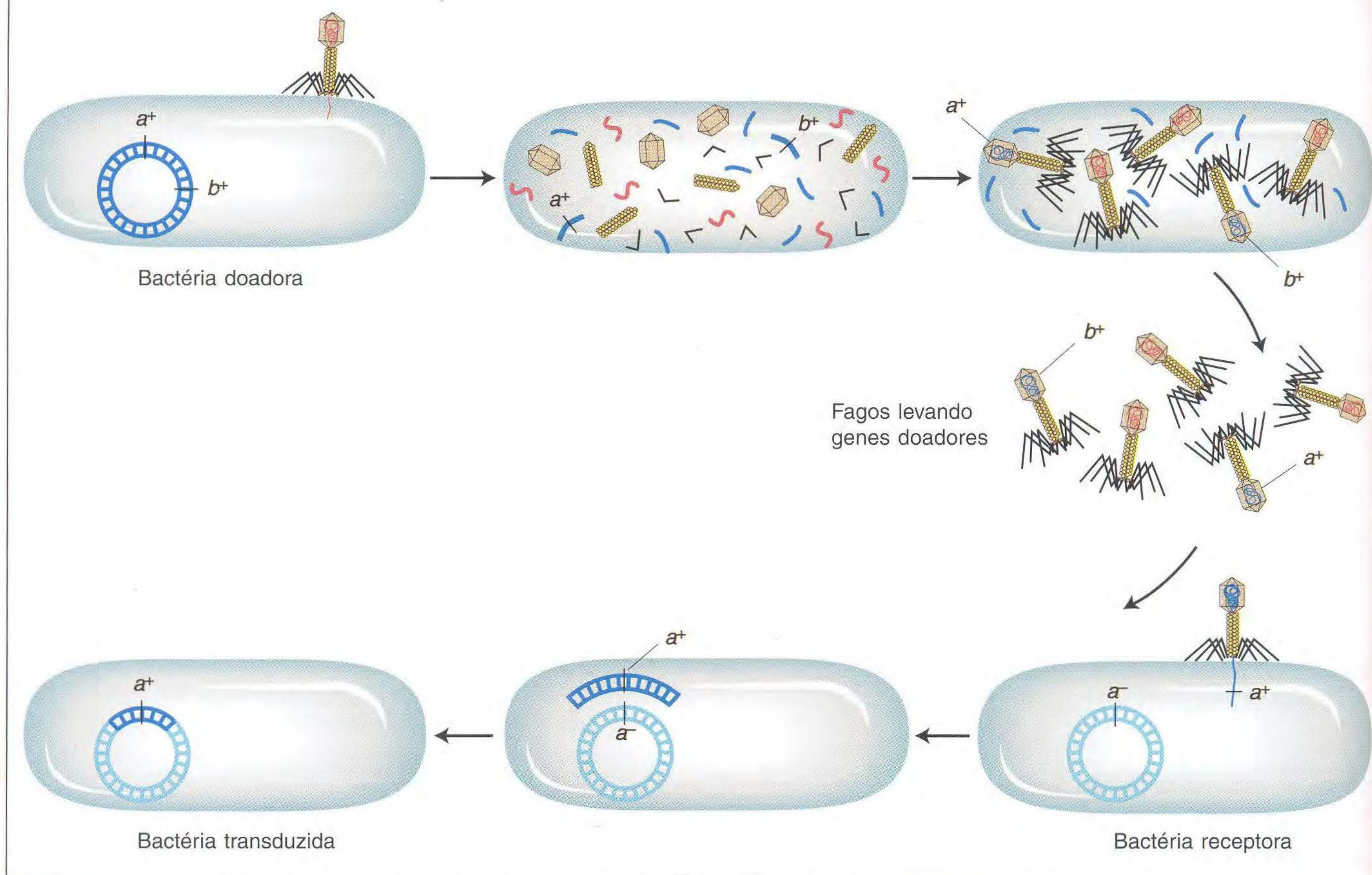
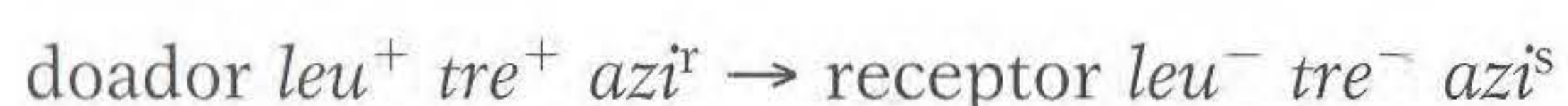


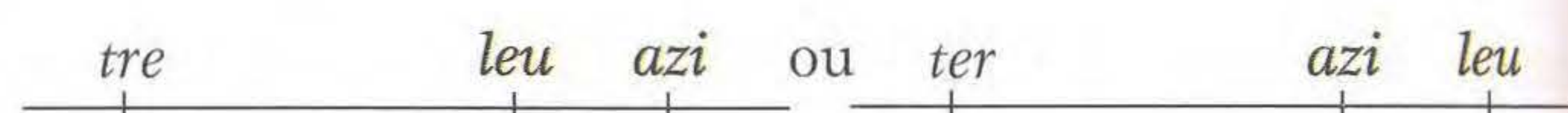
FIG. 5.29 Um fago recém-formado pode captar DNA do cromossomo de sua célula hospedeira (*em cima*) e, então, injetá-lo em uma nova célula (*embaixo, à direita*). O DNA injetado pode inserir-se no novo cromossomo hospedeiro por recombinação (*embaixo, à esquerda*). Na realidade, apenas uma pequena minoria de prole de fagos (1 em 10.000) leva genes doadores.

A transdução generalizada pode ser usada para obter informações sobre ligação bacteriana quando os genes estão tão próximos que o fago pode pegá-los e transferi-los em um único pedaço de DNA. Por exemplo, suponha que queremos encontrar a distância de ligação entre *met* e *arg* em *E. coli*. Podemos cultivar o fago P1 em uma linhagem doadora $met^+ arg^+$ e, então, permitir que os fagos P1 lisem essa linhagem para infectar uma linhagem $met^- arg^-$. Primeiro, é selecionado um alelo doador, digamos met^+ . Então, a percentagem de colônias met^+ que também são arg^+ é dosada. As linhagens transduzidas para ambos, met^+ e arg^+ , são chamadas **co-transdutantes**. Quanto maior a frequência de co-transdução, mais próximos devem estar os dois marcadores (o oposto da maioria das medidas de mapeamento). Os valores de ligação geralmente são expressos como frequências de co-transdução (Figura 5.30).

Usando uma extensão desse enfoque, podemos estimar o tamanho do pedaço do cromossomo hospedeiro que o fago pode captar, como no tipo seguinte de experimento, que usa o fago P1:



Nesse experimento, o fago P1 crescido na linhagem doadora $leu^+ tre^+ azi^r$ infecta a linhagem receptora $leu^- tre^- azi^s$. A estratégia é selecionar um ou mais alelos doadores na receptora e, então, testar essas transdutantes quanto à presença dos alelos não-selecionados. Os resultados são destacados no Quadro 5.3. O experimento 1 no Quadro 5.3 nos diz que *leu* é relativamente próximo de *azi* e distante de *tre*, deixando-nos com duas possibilidades:



O experimento 2 nos diz que *leu* está mais próximo de *tre* do que *azi*, e assim o mapa deve ser



Selecionando tre^+ e leu^+ juntos nos fagos transdutores no experimento 3, vemos que o pedaço transduzido de material genético nunca inclui o locus *azi* porque a cabeça do fago não pode levar um fragmento de DNA tão grande. P1 só pode co-transduzir genes distantes aproximadamente 1,5 min no mapa do cromossomo de *E. coli*.

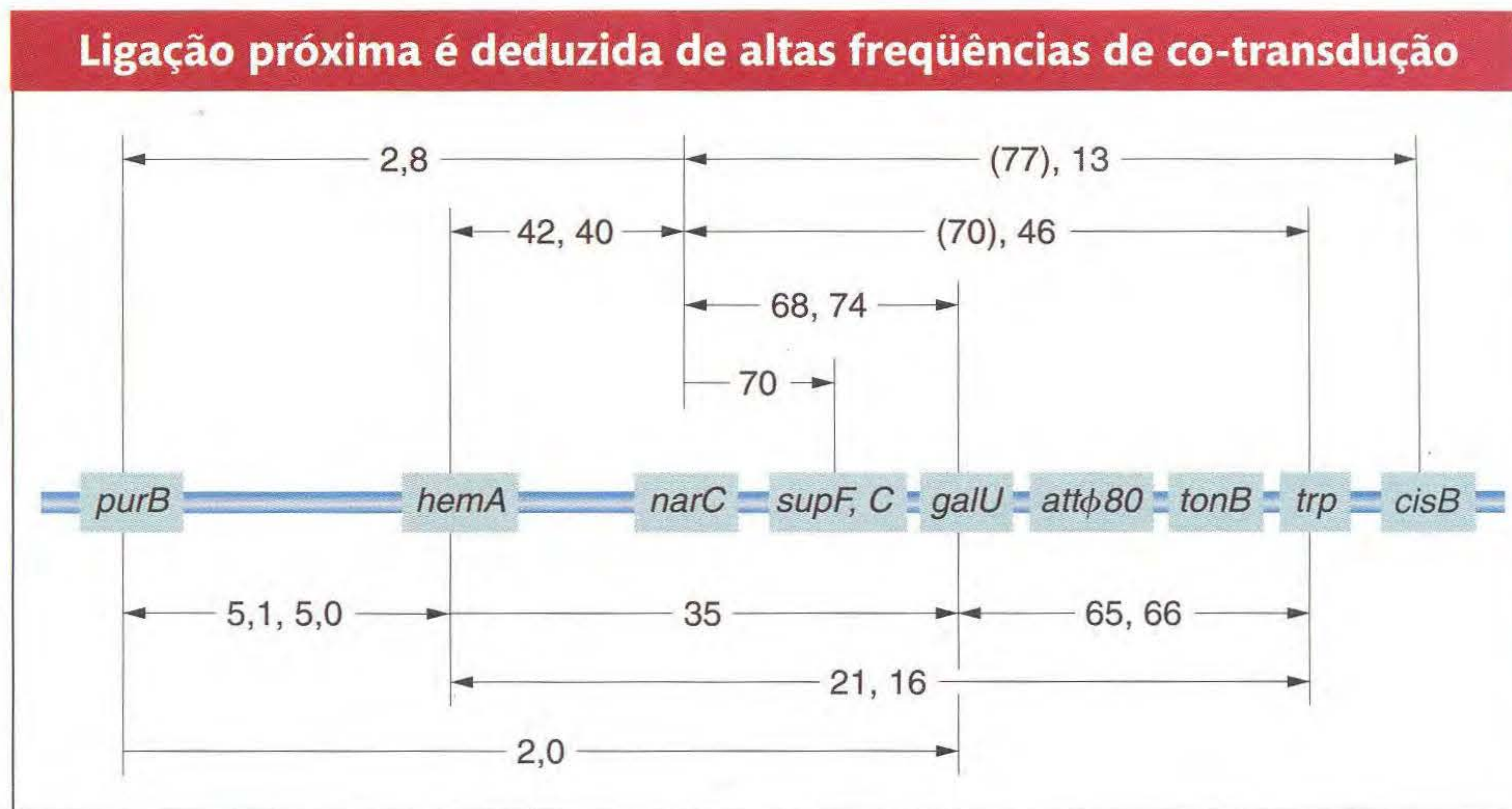


FIG. 5.30 O diagrama mostra um mapa genético da região *purB* até *cysB* de *E. coli* determinada por co-transdução de P1. Os números dados são médias em percentagem das freqüências de co-transdução obtidas em vários experimentos. Os valores entre parênteses são considerados não-confiáveis. [De J.R. Guest, *Mol. Gen. Genet.* 105, 1969, 285.]

Quadro 5.3 Marcadores Acompanhantes em Transduções Específicas de P1

Experimento	Marcador selecionado	Marcadores não-selecionados
1	<i>leu</i> ⁺	50% são <i>azi</i> ^r ; 2% são <i>tre</i> ⁺
2	<i>tre</i> ⁺	3% são <i>leu</i> ⁺ ; 0% são <i>ari</i> ^r
3	<i>leu</i> ⁺ e <i>tre</i> ⁺	0% são <i>ari</i> ^r

Transdução especializada

Um transdutor generalizado, tal como o fago P22, capta aleatoriamente fragmentos do DNA quebrado do hospedeiro. Como outros fagos, que atuam como transdutores especializados, são capazes de levar apenas alguns genes do hospedeiro para as células receptoras? A resposta mais curta é que um transdutor especializado insere-se no cromossomo bacteriano apenas em uma posição. Quando ele sai, ocorre uma alça defeituosa (similar ao tipo que produz os plasmídeos F⁻). Assim, ele pode captar e transduzir apenas genes que estão próximos.

O protótipo da **transdução especializada** foi dado por estudos feitos por Joshua e Esther Lederberg em um fago temperado de *E. coli* chamado *lambda* (λ). O fago λ tornou-se o fago mais intensamente estudado e mais bem caracterizado.

Comportamento do profago O fago λ tem efeitos incomuns quando são usadas células lisogênicas para ele nos cruzamentos. No cruzamento de uma Hfr não-infectada com uma receptora lisogênica F⁻ [Hfr \times F⁻ (λ)], são prontamente recuperados exconjugantes F⁻ lisogênicos com genes Hfr. Entretanto, no cruzamento recíproco Hfr(λ) \times F⁻, os primeiros genes do cromossomo Hfr são recuperados entre os

exconjugantes, mas os recombinantes para genes *tardios* não são recuperados. Além disso, exconjugantes lisogênicos quase nunca são recuperados desse cruzamento recíproco. Qual a explicação? As observações fazem sentido se o profago λ estiver se comportando como um locus gênico bacteriano (isto é, como parte do cromossomo bacteriano). Assim, o profago entraria na célula F em uma época específica correspondendo à sua posição no centrômero. Os genes iniciais são recuperados porque entram antes do profago. Os genes posteriores não são recuperados, pois a lise destrói a célula receptora. Nos experimentos de reprodução interrompida, o profago λ de fato sempre entra na célula F⁻ em uma época específica, proximamente ligado ao locus *gal*.

Em um cruzamento Hfr(λ) \times F⁻, a entrada do profago λ na célula dispara imediatamente o profago em um ciclo lítico; esse processo é chamado de **indução zigótica** (Figura 5.31). Entretanto, no cruzamento de duas células lisogênicas Hfr(λ) \times F⁻ (λ), não há indução zigótica. A presença de qualquer fago impede que outro vírus infectante cause lise. O profago produz um fator citoplasmático que reprime a multiplicação do vírus. (O repressor citoplasmático dirigido por fago explica muito bem a imunidade da bactéria lisogênica, porque um fago imediatamente encontraria um repressor e seria inativado.)

Inserção de λ Os experimentos de reprodução interrompida descritos mostraram que o profago λ é parte de um cromossomo lisogênico da bactéria. Como o profago λ é inserido no genoma bacteriano? Em 1962, Allan Campbell propôs que ele se insere por um único crossing entre um cromossomo circular de fago λ e o cromossomo circular de *E. coli*, como mostrado na Figura 5.32. O ponto do crossing seria entre um sítio específico em λ , o **sítio de ligação**, e um sítio de ligação no cromossomo bacteriano situado entre os genes *gal* e *bio*, porque λ integra-se nessa posição no cromossomo de *E. coli*.

Transferência de profago λ durante conjugação pode disparar lise

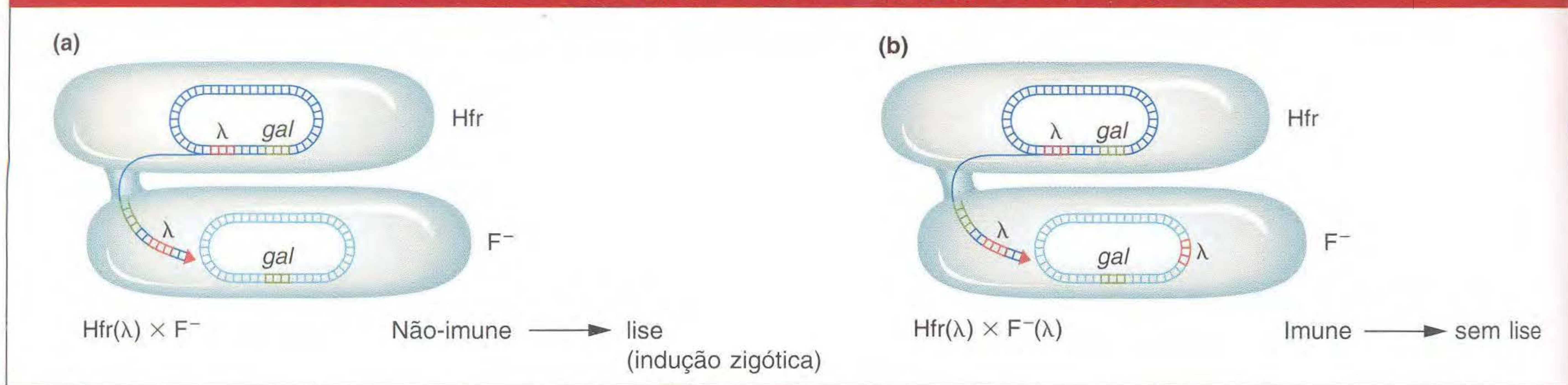


FIG. 5.31 Um profago λ pode ser transferido para uma receptora durante a conjugação, mas o profago dispara a lise, um processo chamado de indução zigótica, apenas se a receptora já não tiver profago, isto é, no caso mostrado na parte a, mas não na parte b.

Um atrativo da proposta de Campbell é que, a partir dela, existem previsões que os geneticistas podem testar. Por exemplo, a integração do profago ao cromossomo de *E. coli* deve aumentar a distância genética entre genes bacterianos flangeadores, como pode ser visto na Figura 5.32 para *gal* e *bio*. De fato, os estudos mostram que a lisogenia *aumenta* o tempo de entrada ou as distâncias de recombinação entre os genes bacterianos. Essa localização única de λ contribui para sua transdução especializada.

Mecanismo de transdução especializada

Como um profago, λ se insere entre a região *gal* e a região *bio* do cromossomo hospedeiro (Figura 5.33), e, nos experimentos de transdução, como esperado, λ só pode transduzir os genes *gal* e *bio*.

Como λ leva genes vizinhos? A explicação está, novamente, em uma reversão imperfeita do mecanismo de inserção de Campbell, como o da formação de F'. O evento de recombinação

entre regiões específicas de λ e o cromossomo bacteriano é catalisado por um sistema enzimático especializado codificado pelo fago que usa o sítio de ligação de λ como um substrato. O sistema enzimático dita que λ integre-se apenas em um ponto específico entre *gal* e *bio* no cromossomo (veja Figura 5.33a). Mais ainda, durante a lise, o profago λ normalmente sai exatamente no ponto para produzir um cromossomo λ circular normal, como visto na Figura 5.33b(i). Muito raramente, a excisão é anormal devido a uma alça anormal. Nesse caso, a alça de DNA do fago pode captar um gene vizinho e deixar alguns genes do fago, como visto na Figura 5.33b(ii). O genoma resultante do fago é defeituoso por causa dos genes deixados, mas também ganhou um gene bacteriano, *gal* ou *bio*. O DNA anormal levando genes vizinhos pode ser embalado nas cabeças dos fagos para produzir partículas de fago que podem infectar outras bactérias. Esses fagos são chamados de λ gal (*gal* sem λ) ou λ dbio. Na presença de um segundo fago normal em uma infecção dupla, o λ gal pode se integrar ao cromossomo no sítio de ligação de λ (Figura 5.33c). Desse modo, os genes *gal* nesse caso são transduzidos para o segundo hospedeiro.

Fago λ se insere por um crossing em sítios específicos

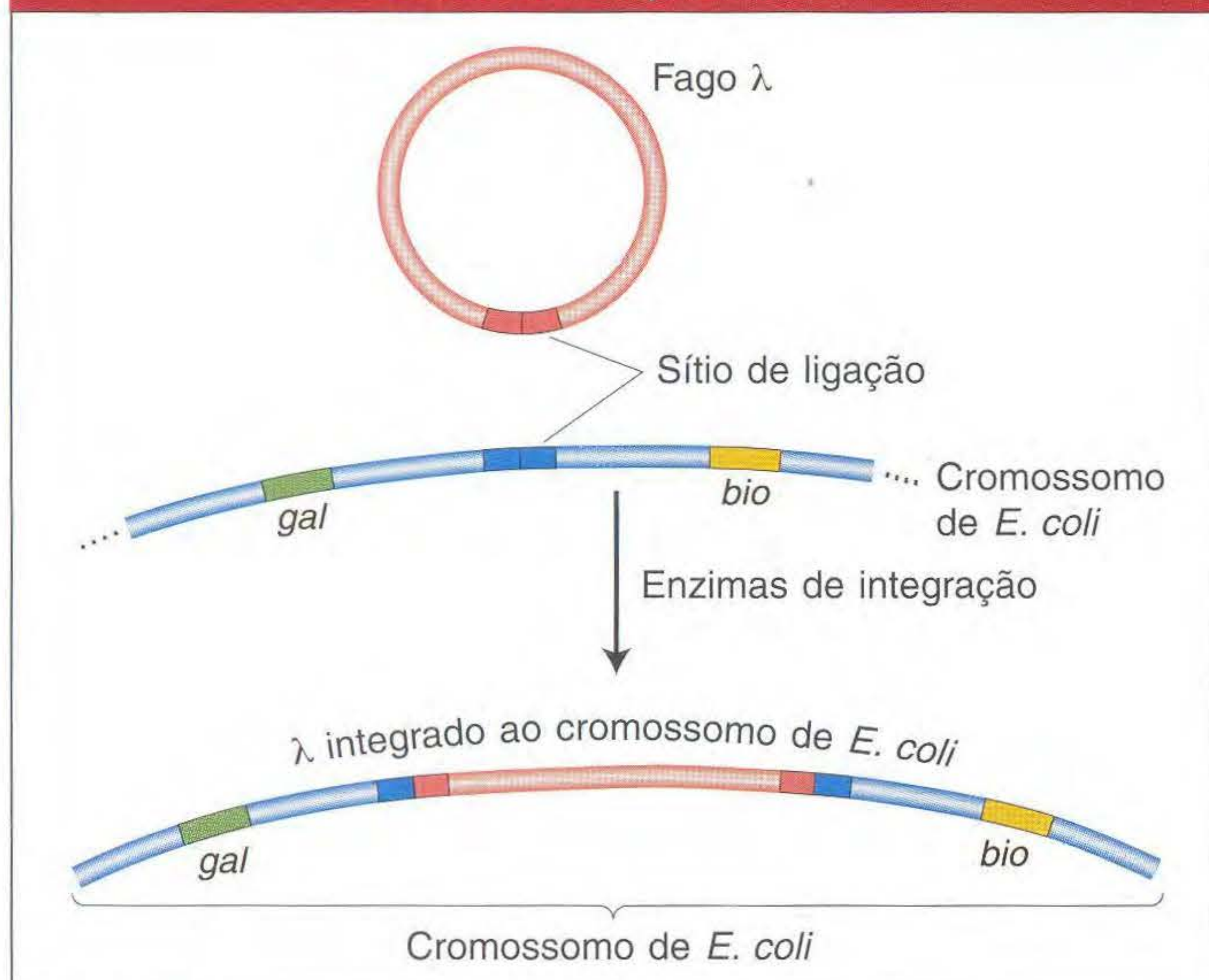


FIG. 5.32 A recombinação recíproca ocorre entre um sítio específico de ligação no DNA λ circular e uma região específica chamada sítio de ligação no cromossomo de *E. coli* entre os genes *gal* e *bio*.

A transdução ocorre quando fagos recém-formados adquirem genes do hospedeiro e os transferem para outras bactérias. A *transdução generalizada* pode transferir qualquer gene do hospedeiro. Ela ocorre quando a embalagem do fago acidentalmente incorpora DNA bacteriano ao invés de DNA do fago. A *transdução especializada* é devida a uma alça defeituosa do profago do cromossomo bacteriano, e, assim, o novo fago inclui tanto o fago quanto os genes bacterianos. O fago transdutor pode transferir apenas genes hospedeiros específicos.

► 5.6 Mapas Físicos e Mapas de Ligação Comparados

Alguns mapas cromossômicos muito detalhados de bactérias foram obtidos combinando as técnicas de mapeamento de reprodução interrompida, mapeamento de recombinação, transformação e transdução. Hoje em dia, novos marcadores

Alça defeituosa produz fago λ contendo DNA bacteriano

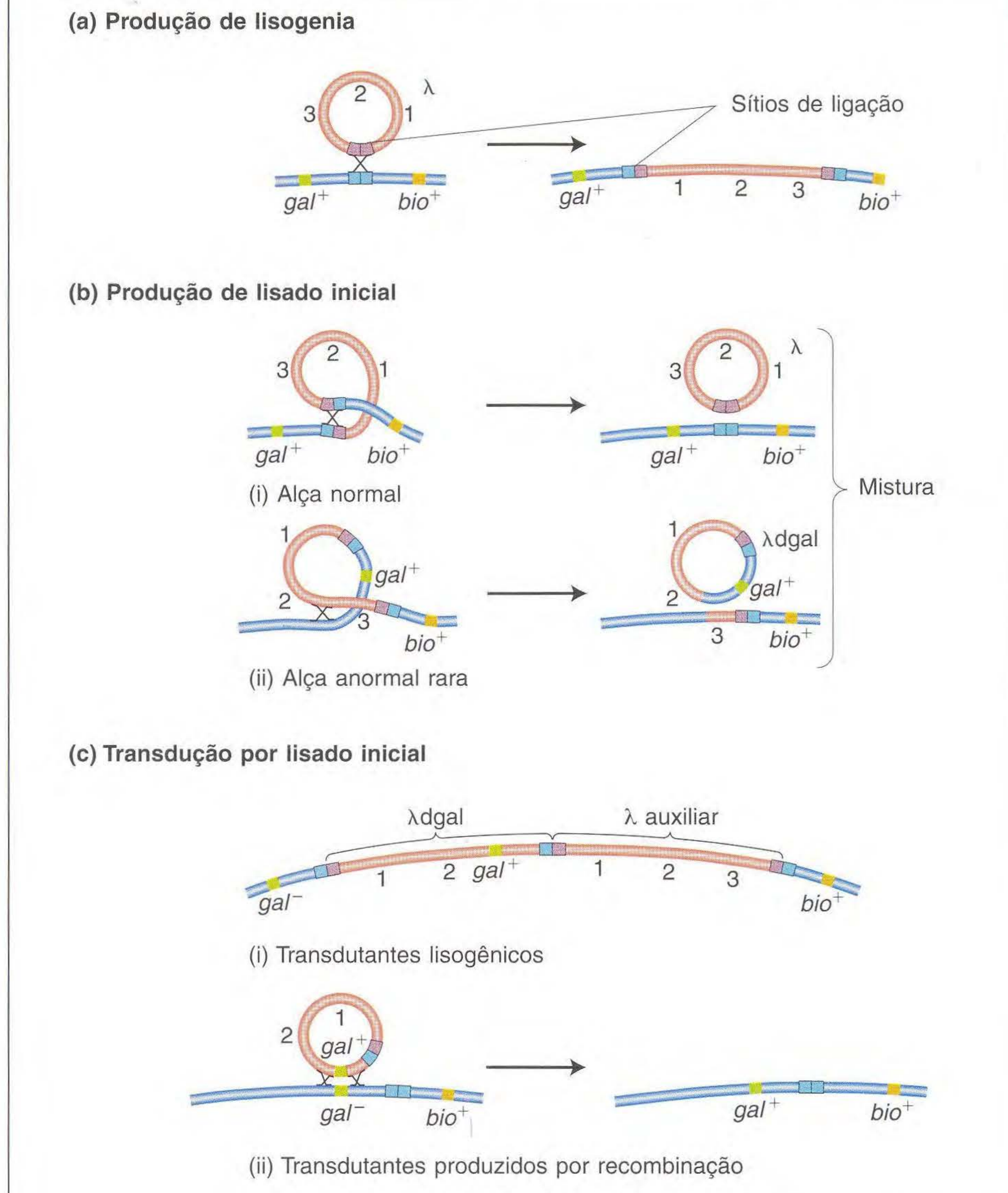


FIG. 5.33 O diagrama mostra como opera a transdução especializada no fago λ . (a) O crossing no sítio especializado de ligação produz uma bactéria lisogênica. (b) A bactéria lisogênica pode produzir λ normal (i) ou, raramente, λ dgal (ii), uma partícula transdutora contendo o gene *gal*. (c) Transdutores *gal*⁺ podem ser produzidos por (i) co-incorporação de λ dgal e λ (agindo como auxiliar) ou (ii) crossings flanqueadores do gene *gal*, um evento raro. Os boxes duplos em azul são sítios de ligação bacteriana, os boxes duplos púrpura são o sítio de ligação de λ e os pares de boxes azul e púrpura são sítios híbridos de integração, derivados parcialmente de *E. coli* e parcialmente de λ .

genéticos são tipicamente mapeados primeiro em um segmento com cerca de 10 a 15 min de mapa usando reprodução interrompida. Marcadores adicionais proximamente ligados podem ser mapeados em uma análise de escala mais fina com o uso de co-transdução de P1 ou recombinação.

Em 1963, o mapa de *E. coli* (Figura 5.34) já detalhou as posições de aproximadamente 100 genes. Após 27 anos de maior refinamento, o mapa de 1990 mostrou as posições de mais de 1.400 genes. A Figura 5.35 mostra um trecho de 5 min do mapa de 1990 (que está ajustado a uma escala de 100 min). A complexidade desses mapas ilustra o poder e sofisticação da análise genética. O quanto esses mapas correspondem à realidade física? Em 1997, a sequência de DNA de todo o genoma de *E. coli* de 4.632.221 pares de bases foi

completado, permitindo-nos comparar a posição exata dos genes no mapa genético com a posição da sequência correspondente codificante na sequência linear do DNA (o mapa físico). O mapa completo é representado na Figura 5.36. A Figura 5.37 faz uma comparação de um segmento de ambos os mapas. Claramente, o mapa genético é uma correspondência próxima do mapa físico.

O Capítulo 4 considera alguns modos pelos quais o mapa físico (geralmente a sequência genômica completa) pode ser útil em mapear novas mutações. Nas bactérias, a técnica de **mutagênese insercional** é outro modo para, rapidamente, enfocar uma mutação posicional em um mapa físico conhecido. A técnica causa mutações pela inserção aleatória de fragmentos de DNA “exógeno”. As inserções inativam qualquer

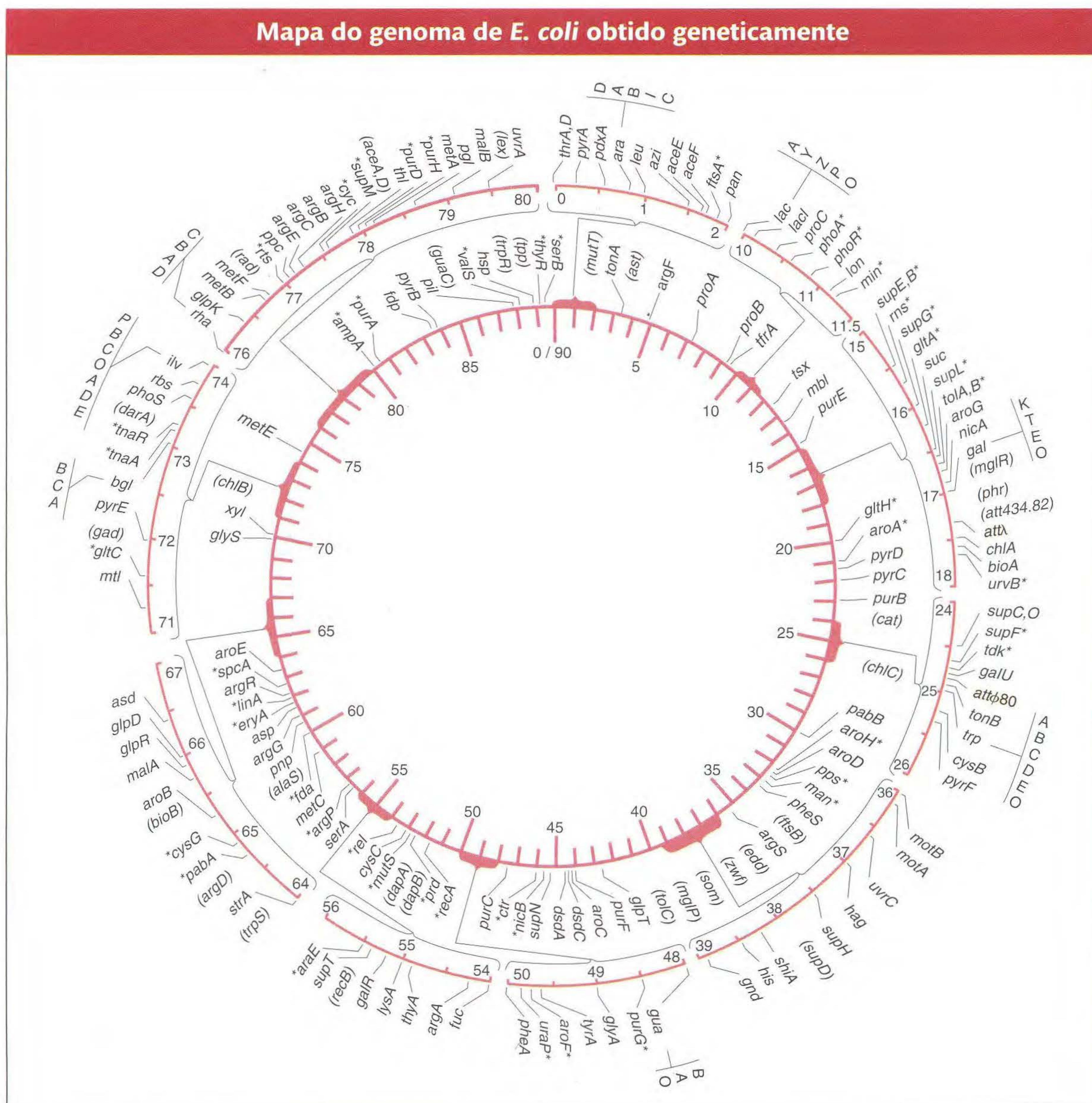


FIG. 5.34 O mapa genético de 1963 dos genes de *E. coli* com fenótipos mutantes. As unidades estão em minutos, com base na reprodução interrompida e experimentos de recombinação. Os asteriscos referem-se a posições de mapa que não são tão precisas quanto as outras posições. [De G.S. Stent, *Molecular Biology of Bacterial Viruses*. Copyright 1963 by W.H. Freeman and Company.]

Parte do mapa físico do genoma de *E. coli*, obtido por sequenciamento

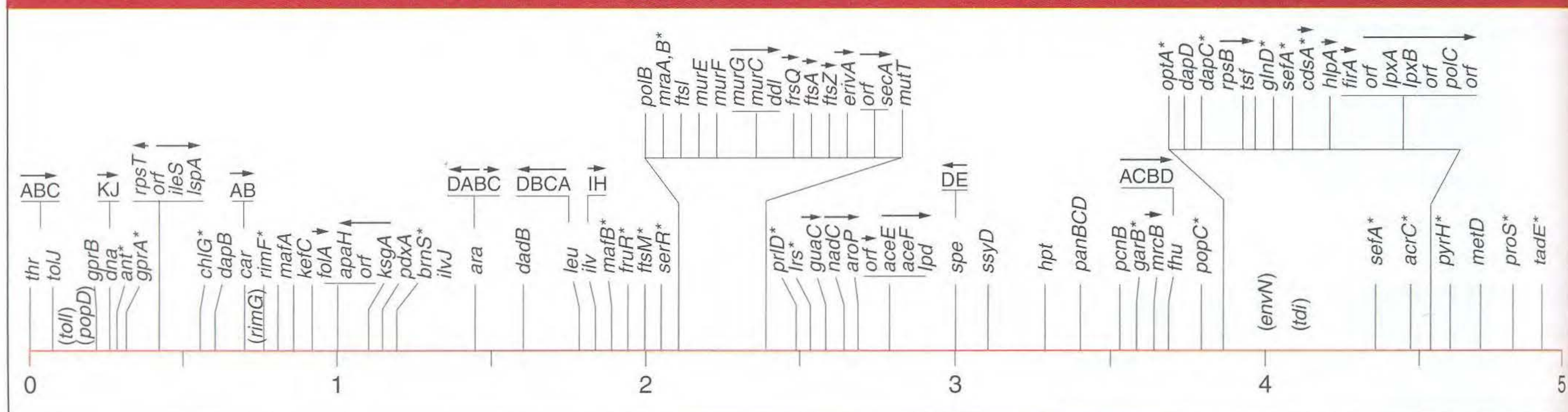


FIG. 5.35 Um desenho em escala linear de uma seção de 5 min do mapa de ligação de *E. coli* de 100 min de 1990. Os parênteses e asteriscos indicam marcadores para os quais a localização exata era desconhecida na época da publicação. As setas acima dos genes e grupos de genes indicam a direção da transcrição. [De B.J. Bachmann, "Linkage Map of *Escherichia coli* K-12. Edition 8", *Microbiol. Rev.* 54, 1990. 130-197.]

Mapa físico do genoma de *E. coli*

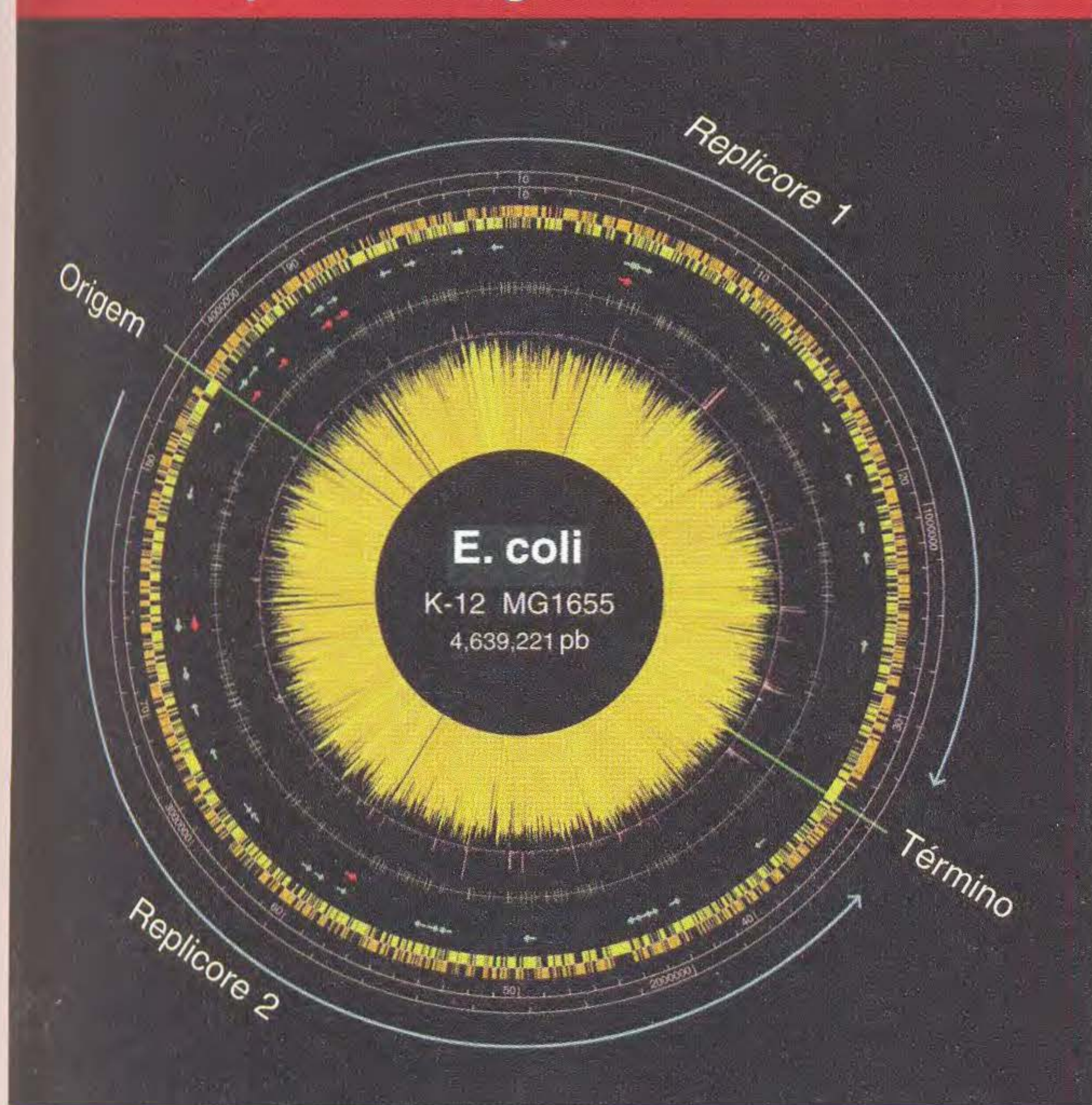


FIG. 5.36 Esse mapa foi obtido do seqüenciamento do DNA e posições de plotagem de genes. Chave dos componentes a partir da parte externa:

- A origem da replicação do DNA e o término estão marcados.
- As duas escalas são em pares de bases de DNA e em minutos.
- Os histogramas em laranja e amarelo mostram a distribuição dos genes nos dois filamentos diferentes de DNA.
- As setas representam genes para rRNA (vermelho) e tRNA (verde)
- A "estrela" central é um histograma de cada gene com linhas de tamanhos que refletem o nível previsto de transcrição.

[De F.R. Blattner et al., "The Complete Genome Sequence of *Escherichia coli* K-12," *Science* 277, 1997, 1453-1462. Imagem por cortesia do Dr. Guy Plunkett III. Reproduzido de AAAS, com permissão.]

gene no qual ocorram interrompendo a unidade transcricional. Os transposons são insertos particularmente úteis para esse fim em vários organismos-modelo, inclusive bactérias. Para mapear uma nova mutação, o procedimento é o seguinte. O DNA de um transposon levando um alelo de resistência ou outro marcador selecionável é introduzido por transformação em receptores bacterianos que não têm transposons ativos. Os transposons inserem-se mais ou menos aleatoriamente, e qualquer um que entre no meio de um gene causa uma mutação. Um subgrupo de todos os mutantes obtidos terá fenótipos relevantes ao processo bacteriano em estudo, e esses fenótipos tornam-se o foco da análise.

A beleza da inserção de transposons é que, como sua seqüência é conhecida, o gene mutante pode ser localizado e seqüenciado. Os primers de replicação do DNA são criados ajustando-se à seqüência conhecida do transposon (veja Capítulo 20). Esses primers são usados para iniciar a análise de seqüenciamento que ocorre *fora* do transposon no gene vizinho. A seqüência curta obtida pode então ser colocada em um computador e comparada com a seqüência genômica completa. Dessa análise, são obtidas a posição do gene e sua seqüência completa. A função de um homólogo desse gene já pode ter sido deduzida em outros organismos. Assim, você pode ver que esse enfoque (como o introduzido no Capítulo 4) é outro modo de unir o fenótipo mutante com a posição de mapa e a potencial função. A Figura 5.38 resume esse enfoque.

Como um adendo, é interessante que muitos dos experimentos históricos que revelam a circularidade dos genomas de bactérias e plasmídeos coincidem com a publicação e popularização do *Senhor dos Anéis* de J.R.R. Tolkien. Consequentemente, uma revisão da genética bacteriana nessa época corresponde à seguinte citação da trilogia:

Um Anel para excluir todos, Um Anel para encontrá-los,
Um Anel para trazer todos e na escuridão ligá-los.

Proporções dos mapas genético e físico são similares, mas não idênticas

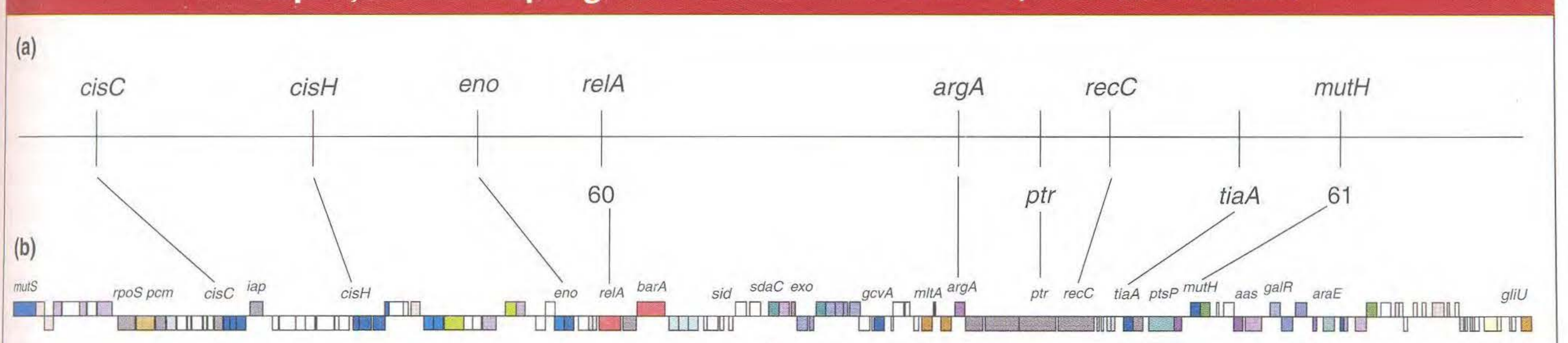


FIG. 5.37 Um alinhamento dos mapas genético e físico. (a) Marcadores no mapa genético de 1990 na região perto de 60 e 61 min. (b) As posições exatas de cada gene, com base na seqüência completa do genoma de *E. coli*. (Note que cada gene é denominado nesse mapa, para simplificação.) Os boxes alongados são genes e genes hipotéticos. Cada cor representa um tipo diferente de função. Por exemplo, vermelho indica funções reguladoras e azul-escuro indica funções na replicação do DNA, recombinação e reparo. As linhas entre os mapas nas partes a e b conectam o mesmo gene em cada mapa. [De F.R. Blattner et al., "The Complete Genome Sequence of *Escherichia coli* K-12," *Science*, vol. 277, 5 de setembro de 1997, pp. 1453-1462. Imagem por cortesia do Dr. Guy Plunkett III.]

Mutagênese por transposon pode ser usada para mapear uma mutação na sequência do genoma

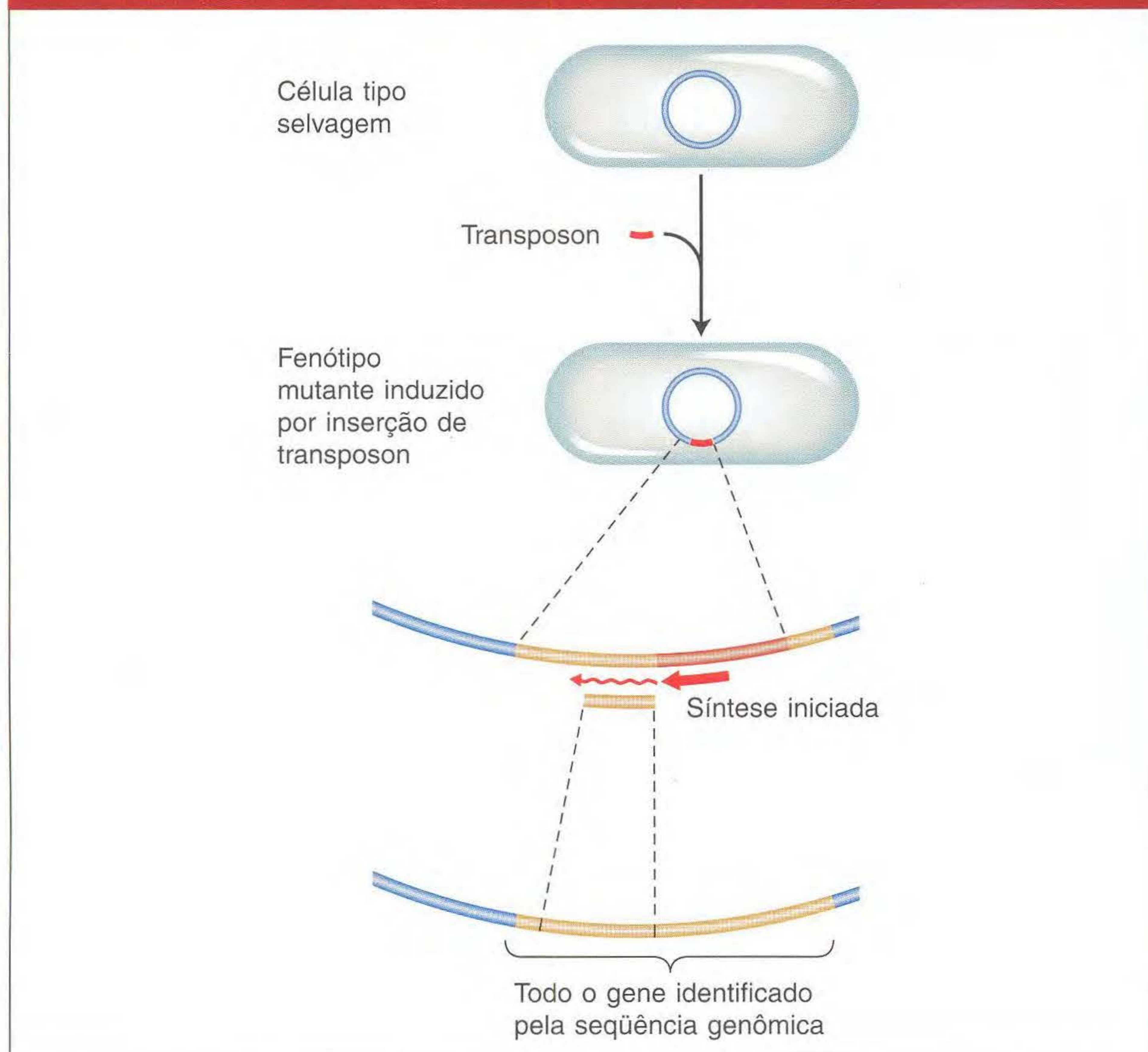


FIG. 5.38 A inserção de um transposon insere uma mutação em um gene de posição e função desconhecidos. O segmento próximo ao transposon é replicado, sequenciado e pareado a um segmento na sequência genômica completa.

► Resumo

Os avanços na genética de bactérias e fagos nos últimos 50 anos forneceram os fundamentos para a biologia molecular e a clonagem (discutida em capítulos posteriores). No começo desse período, a transferência gênica e a recombinação foram encontradas ocorrendo entre linhagens diferentes de bactérias. Nas bactérias, entretanto, o material genético é passado em uma só direção — por exemplo, na *Escherichia coli*, de uma célula doadora (F^+ ou Hfr) para uma célula receptora (F^-). A habilidade do doador é determinada pela presença na célula de um fator de fertilidade (F), um tipo de plasmídeo. Em certas ocasiões, o fator F presente no estado livre nas células F^+ pode se integrar ao cromossomo de *E. coli* e formar uma célula Hfr. Quando isso ocorre, um fragmento do cromossomo doador pode transferir-se para uma célula receptora e, subsequentemente, recombinar-se com o cromossomo receptor. Como o fator F pode se inserir em locais diferentes do cromossomo hospedeiro, os primeiros pesquisadores foram capazes de juntar os fragmentos transferidos para mostrar que o cromossomo de *E. coli* é um único círculo, ou anel. A interrupção da transferência em tempos diferentes deu aos geneticistas um método não-convencional (reprodução interrompida) para construir um

mapa de ligação do único cromossomo de *E. coli* e outras bactérias similares, nas quais a unidade de mapa é uma unidade de tempo (minutos). Em uma extensão dessa técnica, a frequência de recombinantes entre marcadores conhecidos como tendo entrado na receptora pode fornecer uma distância de mapa em escala fina.

Vários tipos de plasmídeos além de F podem ser encontrados. Os plasmídeos R levam alelos de resistência a antibióticos, em geral dentro de um elemento móvel chamado transposon. A rápida dispersão do plasmídeo causa resistência de toda uma população a medicamentos importantes. Os derivados de tais plasmídeos naturais tornaram-se vetores de clonagem importantes, úteis para o isolamento gênico e estudo em todos os organismos.

As características genéticas também podem ser transferidas de uma bactéria para outra sob a forma de pedaços de DNA levados para a célula do ambiente extracelular. Esse processo de transformação nas bactérias foi a primeira demonstração de que o DNA é o material genético. Para que ocorra a transformação, o DNA deve ser levado para uma célula receptora, e a recombinação deve ocorrer entre um cromossomo receptor e o DNA incorporado.

As bactérias podem ser infectadas por vírus chamados bacteriófagos. Em um método de infecção, o cromossomo do fago pode entrar na célula bacteriana e, usando a maquinaria metabólica bacteriana, produzir uma prole de fagos que rompem a bactéria hospedeira. Os novos fagos podem então infectar outras células. Se dois fagos de genótipos diferentes infectam a mesma hospedeira, pode ocorrer recombinação entre seus cromossomos.

Em outro modo de infecção, a lisogenia, o fago injetado fica dormente na bactéria. Em muitos casos, esse fago dormente (o profago) incorpora-se ao cromossomo hospedeiro e replica-se com ele. Seja espontaneamente ou sob estimulação apropriada, o profago pode deixar seu estado dormente e lisar a bactéria hospedeira.

Um fago pode levar genes bacterianos do doador para um receptor. Na transdução generalizada, o DNA hospedeiro aleatório é incorporado sozinho na cabeça do fago durante a lise. Na transdução especializada, a excisão defeituosa do profago de um único locus cromossômico resulta na inclusão de genes específicos do hospedeiro bem como DNA do fago na cabeça do fago.

Hoje em dia, um mapa físico sob a forma da sequência genômica completa está disponível para muitas espécies de bactérias. Com o uso desse mapa físico genômico, a posição de mapa de uma mutação de interesse pode ser localizada com precisão. Primeiro, são produzidas mutações apropriadas pela inserção de transposons (mutagênese insercional). Então, a sequência de DNA que circunda o transposon inserido é obtida e correlacionada a uma sequência no mapa físico. Essa técnica fornece o locus, a sequência e, possivelmente, a função do gene de interesse.

Termos-chave

auxotrófico	indução zigótica	profago
bactéria lisogênica	infecção dupla (mista)	prototrófica
bacteriófago (fago)	infecção mista (dupla)	receptora
clone celular	lisado	recombinação de fago
colônia	lise	replicação em círculo rolante
conjugação	marcador genético	reprodução interrompida
co-transdutante	marcador não-selecionado	sistema seletivo
doador	meio mínimo	sítio de ligação
endogenoto	merozigoto	término
exconjugante	mutagênese insercional	transdução
exogenoto	mutante resistente	transdução especializada
F ⁺ (doadora)	origem (O)	transdução generalizada
F ⁻ (receptora)	placa	transformação
fago (bacteriófago)	plaqueamento	transformação dupla
fago temperado	plasmídeo	triagem
fago virulento	plasmídeo F'	vírus
fator de fertilidade (F)	plasmídeo R	
Hfr (alta frequência de recombinação)	procarionte	

Problemas Resolvidos

Problema resolvido 1. Suponha que uma célula fosse incapaz de fazer recombinação generalizada (*rec*⁻). Como essa célula se comportaria como receptora na transdução generalizada e na especializada? Primeiro, compare cada tipo de transdução e, então, determine o efeito da mutação *rec*⁻ na herança de genes por cada processo.

SOLUÇÃO

A transdução generalizada envolve a incorporação de fragmentos cromossômicos nas cabeças dos fagos, que então infectam linhagens receptoras. Os fragmentos do cromossomo são incorporados aleatoriamente nas cabeças dos fagos, e, assim, qualquer marcador no cromossomo hospedeiro bacteriano pode ser transduzido para outra linhagem por transdução generalizada. Em contraste, a transdução generalizada envolve a integração do fago em um ponto específico do cromossomo, e a rara incorporação dos marcadores cro-

mossômicos perto do sítio de integração no genoma do fago. Portanto, apenas esses marcadores que estão perto do sítio específico de integração do fago ao cromossomo hospedeiro podem ser transduzidos.

Os marcadores são herdados por vias diferentes na transdução generalizada e especializada. Um fago de transdução generalizada injeta um fragmento do cromossomo doador no receptor. Esse fragmento deve ser incorporado ao cromossomo receptor por recombinação, com o uso do sistema de recombinação do receptor. Portanto, um receptor *rec*⁻ não será capaz de incorporar fragmentos de DNA e não pode herdar marcadores por transdução generalizada. Por outro lado, a via principal de herança dos marcadores por transdução especializada é pela integração da partícula de transdução especializada no cromossomo hospedeiro no sítio específico de integração do fago. Essa integração, que às vezes requer um fago tipo selvagem adicional (auxiliar), é mediada por

um sistema enzimático específico do fago que é independente das enzimas normais de recombinação. Portanto, uma receptora rec^- pode ainda herdar marcadores genéticos por transdução especializada.

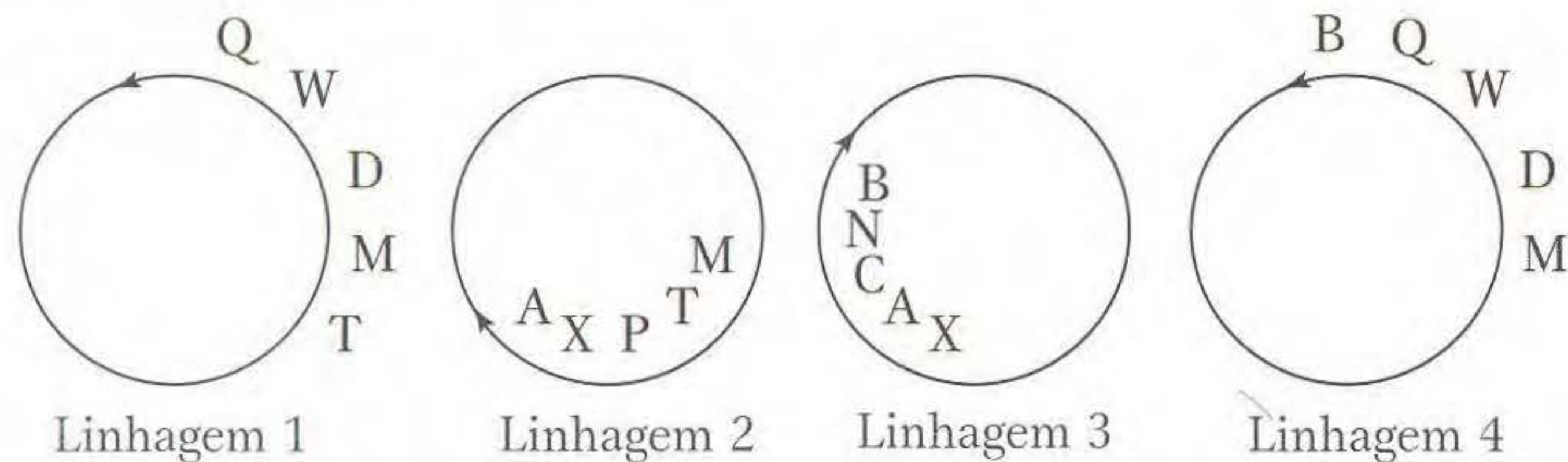
Problema resolvido 2. Em *E. coli*, quatro linhagens Hfr doam os seguintes marcadores genéticos, mostrados na ordem doada:

Linhagem 1:	Q	W	D	M	T
Linhagem 2:	A	X	P	T	M
Linhagem 3:	B	N	C	A	X
Linhagem 4:	B	Q	W	D	M

Todas essas linhagens Hfr são derivadas da mesma linhagem F^+ . Qual a ordem desses marcadores no cromossomo circular da F^+ original?

SOLUÇÃO

Um enfoque de duas etapas funciona bem: (1) determine o princípio subjacente e (2) desenhe um diagrama. Aqui, o princípio é claramente que cada linhagem Hfr doa marcadores genéticos a partir de um ponto fixo no cromossomo circular e que os primeiros marcadores são doados com a mais alta frequência. Como nem todos os marcadores são doados por cada Hfr, apenas os marcadores iniciais devem ser doados para cada Hfr. Cada linhagem nos permite desenhar os seguintes círculos:



Por essa informação, podemos consolidar cada círculo em um mapa de ligação circular da ordem Q, W, D, M, T, P, X, A, C, N, B, Q.

Problema resolvido 3. Em um cruzamento $Hfr \times F^-$, leu^+ entra como o primeiro marcador, mas a ordem dos outros marcadores é desconhecida. Se a Hfr é tipo selvagem e a F^- é auxotrófica para cada marcador em questão, qual a ordem dos marcadores em um cruzamento onde os recombinantes leu^+ são selecionados se 27% são ile^+ , 13% são mal^+ , 82% são tre^+ e 1% são trp^+ ?

SOLUÇÃO

Lembre que a quebra espontânea cria um gradiente natural de transferência, o que torna cada vez menos provável uma receptora receber marcadores cada vez mais posteriores. Como nós selecionamos quanto ao primeiro marcador nesse cruzamento, a frequência de recombinantes é uma função da ordem de entrada para cada marcador. Portanto, podemos determinar imediatamente a ordem dos marcadores genéticos simplesmente olhando as percentagens de recombinantes para qualquer marcador entre os leu^+ recombinantes. Como a herança

de tre^+ é a mais alta, tre^+ deve ser o primeiro marcador a entrar depois de leu . A ordem completa é leu, tre, ile, mal, trp .

Problema resolvido 4. É feito um cruzamento entre uma Hfr, que é $met^+ tia^+ pur^+$, e uma F^- , que é $met^+ tia^- pur^-$. Os estudos de reprodução interrompida mostram que met^+ entra na receptora por último, e, portanto, os recombinantes met^+ são selecionados em um meio contendo suplementos que satisfazem apenas os requisitos pur e tia . Esses recombinantes são testados quanto à presença dos alelos tia^+ e pur^+ . Os seguintes números de indivíduos são encontrados em cada genótipo:

$met^+ tia^+ pur^+$	280
$met^+ tia^+ pur^-$	0
$met^+ tia^- pur^+$	6
$met^+ tia^- pur^-$	52

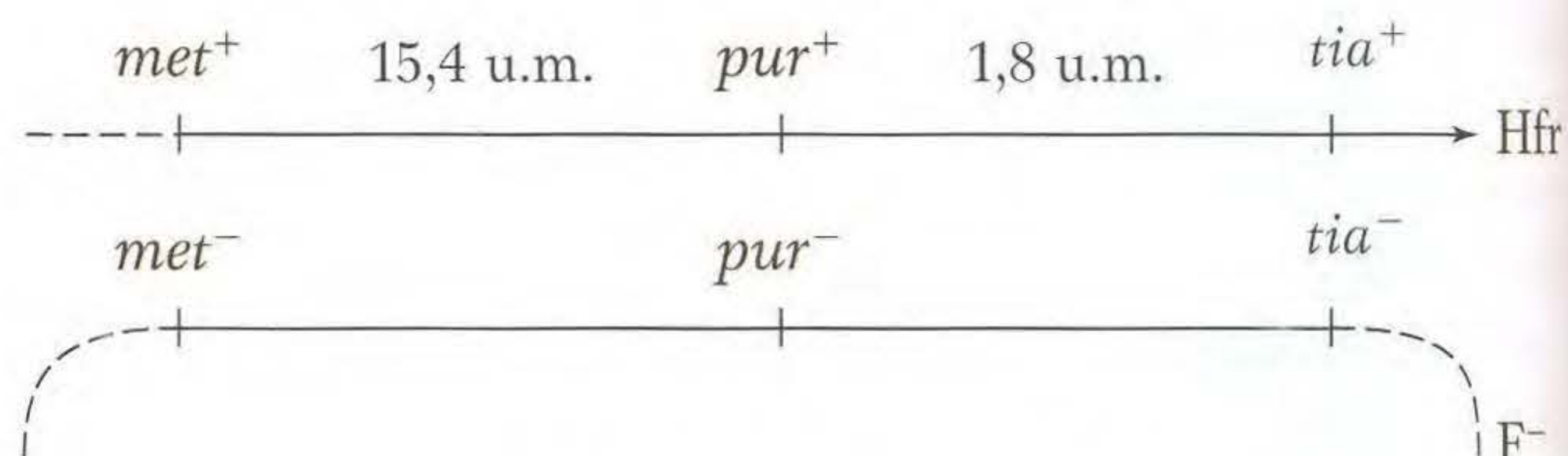
- Por que metionina (Met) foi deixada fora do meio de seleção?
- Qual é a ordem genética?
- Quais as distâncias de mapa nas unidades recombinantes?

SOLUÇÃO

a. Metionina foi deixada fora do meio para permitir a seleção de recombinantes met^+ , pois met^+ é o último marcador a entrar na receptora. A seleção para met^+ garante que todos os loci que são considerados no cruzamento terão já entrado em cada recombinante que analisamos.

b. Aqui, um diagrama das possíveis ordens dos genes é útil. Como sabemos que met entra na receptora por último, existem apenas duas possíveis ordens de genes se o primeiro marcador entrar à direita: met, tia, pur ou met, pur, tia . Como podemos distinguir entre essas duas ordens? Felizmente, uma das quatro classes possíveis de recombinantes requer dois crossings adicionais. Cada ordem possível prevê uma classe diferente que surge por quatro crossings em vez de dois. Por exemplo, se a ordem fosse met, tia, pur , então os recombinantes $met^+ tia^- pur^+$ seriam muito raros. Por outro lado, se a ordem fosse met, pur, tia , então a classe de quatro crossings seria $met^+ pur^- tia^+$. Pelas informações dadas no quadro, $met^+ pur^- tia^+$ é claramente a classe de quatro crossings, e, portanto, a ordem dos genes met, pur, tia é correta.

c. Observe o seguinte diagrama:



Para computarmos a distância entre met e pur , computamos a percentagem de $met^+ pur^- tia^-$, que é $52/388 = 15,4$ u.m. Similarmente, a distância entre pur e tia é $6/338 = 1,8$ u.m.

Problema resolvido 5. Compare o mecanismo de transferência e herança dos genes lac^+ nos cruzamentos com

as linhagens Hfr, F^+ e $F' lac^+$. Como uma célula F^- que não pode sofrer recombinação homóloga normal (rec^-) comporta-se em cruzamentos com cada uma dessas três linhagens? A célula seria capaz de herdar o gene lac^+ ?

SOLUÇÃO

Cada uma dessas três linhagens doa genes por conjugação. Nas linhagens Hfr e F^+ , os genes lac^+ no cromossomo hospedeiro são doados. Na linhagem Hfr, o fator F é integrado ao cromossomo em cada célula, e, assim, os marcadores cromossômicos podem ser doados eficientemente, particularmente se um marcador está perto do sítio de integração de F e é **doado inicialmente**. A população de células F^+ contém uma pequena percentagem de células Hfr, na qual F é integrada ao cromossomo. Essas células são responsáveis pela transferência gênica apresentada pelas culturas de células F^+ . Na transferência de genes mediada por Hfr e F^+ , a herança requer a

incorporação de um fragmento transferido por recombinação (lembre que são necessários dois crossings) no cromossomo F^- . Portanto, uma linhagem F^- que não pode sofrer recombinação não pode herdar marcadores cromossômicos doadores, muito embora sejam transferidos por linhagens Hfr ou células Hfr nas linhagens F^+ . O fragmento não pode ser incorporado ao cromossomo por recombinação. Como esses fragmentos não possuem a habilidade de se replicar dentro da célula F^- , eles são rapidamente diluídos durante a divisão celular.

Ao contrário das células Hfr, as células F' transferem genes levados no fator F' , um processo que não requer a transferência de cromossomos. Nesse caso, os genes lac^+ são ligados ao fator F' e são transferidos com ele com alta eficiência. Na célula F^- , não é necessária nenhuma recombinação, pois a linhagem $F' lac^+$ pode se replicar e ser mantida na população de células em divisão F^- . Portanto, os genes lac^+ são herdados mesmo em uma linhagem rec^- .

Problemas

PROBLEMAS BÁSICOS

1. Descreva o estado do fator F nas linhagens Hfr, F^+ e F^- .
2. Como uma cultura de células F^+ transfere marcadores do cromossomo hospedeiro para uma receptora?
3. Com relação à transferência gênica e integração do gene transferido para um genoma receptor, compare
 - a. Hfr cruza por conjugação e transdução generalizada.
 - b. Derivados de F' , tais como $F' lac$ e transdução especializada.
4. Por que a transdução generalizada é capaz de transferir qualquer gene, mas a transdução especializada é restrita apenas a um pequeno grupo?
5. Uma geneticista microbiana isola uma nova mutação em *E. coli* e quer mapear sua localização cromossômica. Ela usa experimentos de reprodução interrompida com linhagens Hfr e experimentos de transdução generalizada com fago P1. Explique por que cada técnica, em si, é insuficiente para um mapeamento apurado.
6. Em *E. coli*, quatro linhagens Hfr doam os seguintes marcadores, mostrados na ordem doada:

Linhagem 1:	M	Z	X	W	C
Linhagem 2:	L	A	N	C	W
Linhagem 3:	A	L	B	R	U
Linhagem 4:	Z	M	U	R	B

- Todas essas linhagens Hfr são derivadas da mesma linhagem F^+ . Qual a ordem desses marcadores no cromossomo circular da F^+ original?
7. Você recebe duas linhagens de *E. coli*. A linhagem Hfr é $arg^+ ala^+ glu^+ pro^+ leu^+ T^s$; a linhagem F^- é $arg^- ala^- glu^- pro^- leu^- T^r$. Todos os marcadores são nutricionais, exceto T , que determina sensibilidade ou resistência ao fago T1. A ordem de entrada é como a citada, com arg^+ entrando na receptora primeiro e T^s por último. Você descobre que a linhagem F^- morre quando exposta à penicilina (pen^s), mas a linhagem Hfr não (pen^r). Como

você situa o locus para *pen* no cromossomo bacteriano com relação a *arg*, *ala*, *glu*, *pro* e *leu*? Formule sua resposta em etapas lógicas, bem explicadas, e desenhe diagramas explícitos quando possível.



8. É feito um cruzamento entre duas linhagens de *E. coli*: $Hfr arg^+ bio^+ leu^+ \times F^- arg^- bio^- leu^-$. Os estudos de reprodução interrompida mostram que arg^+ entra na receptora por último, e, assim, os recombinantes arg^+ são selecionados em um meio contendo só *bio* e *leu*. Esses recombinantes são testados quanto à presença de bio^+ e leu^+ . Os números seguintes de indivíduos são encontrados para cada genótipo:

$arg^+ bio^+ leu^+$	320	$arg^+ bio^- leu^+$	0
$arg^+ bio^+ leu^-$	8	$arg^+ bio^- leu^-$	48

- a. Qual a ordem dos genes?
 - b. Quais as distâncias de mapa nas percentagens de recombinação?
9. Os mapas de ligação em uma linhagem bacteriana Hfr são calculados em unidades de minutos (o número de minutos entre os genes indica o tempo que leva para o segundo gene seguir o primeiro na conjugação). Ao fazer tais mapas, o geneticista microbiano supõe que o cromossomo bacteriano é transferido de Hfr para F^- a uma velocidade constante. Assim, dois genes separados por 10 min perto da ponta original são supostos como tendo a mesma distância física que dois genes separados por 10 min perto da ponta de ligação de F^- . Sugira um experimento crítico para testar a validade dessa suposição.
 10. Determinada linhagem de Hfr normalmente transmite o marcador pro^+ como o último na conjugação. Em um cruzamento dessa linhagem com uma linhagem F^- , alguns recombinantes pro^+ são recuperados no início do processo de reprodução. Quando essas células pro^+ são misturadas com células F^- , a maioria das células F^- são

convertidas em células *pro*⁺ que também levam o fator F. Explique esses resultados.

11. As linhagens F' em *E. coli* são derivadas de linhagens Hfr. Em alguns casos, essas linhagens F' mostram uma alta taxa de integração com o cromossomo bacteriano de uma segunda linhagem. Além disso, o sítio de integração geralmente é o sítio ocupado pelo fator sexual na Hfr original (antes da produção de linhagens F'). Explique esses resultados.
12. Você tem duas linhagens de *E. coli*, F⁻ *str*^r *ala*⁻ e Hfr *str*^s *ala*⁺, na qual o fator F é inserido perto de *ala*⁺. Crie um teste de triagem para detectar linhagens portadoras de F' *ala*⁺.
13. Cinco linhagens Hfr de A a E são derivadas de uma única linhagem F⁺ de *E. coli*. O quadro seguinte mostra os tempos de entrada dos primeiros cinco marcadores em uma linhagem F⁻ quando cada uma é usada em um experimento de conjugação interrompida:

A	B	C	D	E
<i>mal</i> ⁺ (1)	<i>ade</i> ⁺ (13)	<i>pro</i> ⁺ (3)	<i>pro</i> ⁺ (10)	<i>his</i> ⁺ (7)
<i>str</i> ^s (11)	<i>his</i> ⁺ (28)	<i>met</i> ⁺ (29)	<i>gal</i> ⁺ (16)	<i>gal</i> ⁺ (17)
<i>ser</i> ⁺ (16)	<i>gal</i> ⁺ (38)	<i>xyl</i> ⁺ (32)	<i>his</i> ⁺ (26)	<i>pro</i> ⁺ (23)
<i>ade</i> ⁺ (36)	<i>pro</i> ⁺ (44)	<i>mal</i> ⁺ (37)	<i>ade</i> ⁺ (41)	<i>met</i> ⁺ (49)
<i>his</i> ⁺ (51)	<i>met</i> ⁺ (70)	<i>str</i> ^s (47)	<i>ser</i> ⁺ (61)	<i>xyl</i> ⁺ (52)

- a. Desenhe um mapa da linhagem F⁺, indicando as posições de todos os genes e suas distâncias em minutos.
- b. Mostre o ponto de inserção e orientação do plasmídeo F em cada linhagem Hfr.
- c. No uso de cada uma dessas linhagens Hfr, diga qual alelo você escolheria para obter a mais alta proporção de Hfr exconjugantes.
14. As células de *Streptococcus pneumoniae* com o genótipo *str*^s *mtl*⁻ são transformadas pelo DNA doador de genótipo *str*^r *mtl*⁺ e (em um experimento separado) por uma mistura de dois DNA com genótipos *str*^r *mtl*⁻ e *str*^s *mtl*⁺. O quadro seguinte mostra os resultados.

DNA transformante	Percentagem de células transformadas em		
	<i>str</i> ^r <i>mtl</i> ⁻	<i>str</i> ^s <i>mtl</i> ⁺	<i>str</i> ^r <i>mtl</i> ⁺
<i>str</i> ^r <i>mtl</i> ⁺	4,3	0,40	0,17
<i>str</i> ^r <i>mtl</i> ⁻ + <i>str</i> ^s <i>mtl</i> ⁺	2,8	0,85	0,0066

- a. O que a primeira fila do quadro lhe diz? Por quê?
- b. O que a segunda fila do quadro lhe diz? Por quê?
15. Lembre que, no Capítulo 4, consideramos a possibilidade de que um evento de crossing possa afetar a probabilidade de outro crossing. No bacteriófago T4, o gene *a* dista 1,0 u.m. do gene *b*, que dista 0,2 u.m. do gene *c*. A ordem dos genes é *a, b, c*. Em um experimento de recombinação, você recupera cinco crossings duplos entre *a* e *c* de uma prole de 100.000 vírus. É correto concluir que a interferência é negativa? Explique sua resposta.
16. Você infectou células de *E. coli* com duas linhagens do vírus T4. Uma linhagem é diminuta (*m*), lise rápida (*r*) e turva (*tu*); a outra é tipo selvagem para todos os três mar-

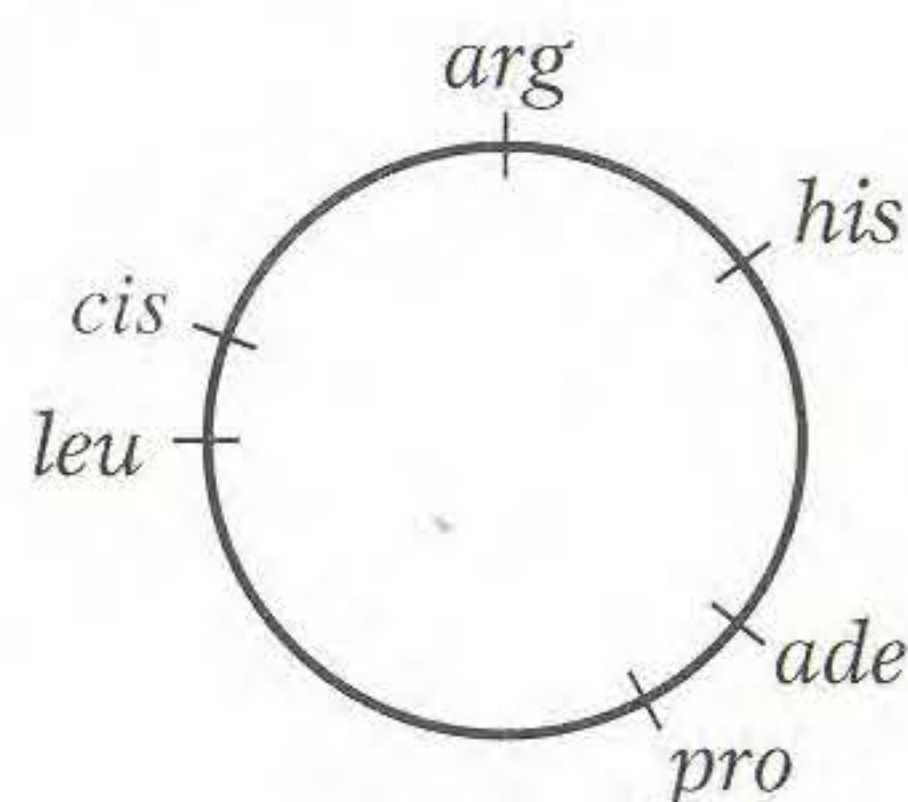
cadores. Os produtos líticos dessa infecção são plaqueados e classificados. As 10.342 placas resultantes foram distribuídas entre oito genótipos, do seguinte modo:

<i>m r tu</i>	3.467	<i>m + +</i>	520
<i>+ + +</i>	3.729	<i>+ r tu</i>	474
<i>m r +</i>	853	<i>+ r +</i>	172
<i>m + tu</i>	162	<i>+ + tu</i>	965

- a. Determine as distâncias de ligação entre *m* e *r*, entre *r* e *tu* e entre *m* e *tu*.
- b. Que ordem de ligação você sugere para os três genes?
- c. Qual o coeficiente de coincidência (veja Capítulo 4) nesse cruzamento? O que ele significa?
- (O problema 16 foi reimpresso com permissão de Macmillan Publishing Co., Inc., de Monroe W. Strickberger, *Genetics*. Copyright 1968 by Monroe W. Strickberger.)
17. Com o uso de P22 como um fago transdutor generalizado cultivado em uma bactéria doadora *pur*⁺ *pro*⁺ *his*⁺, uma linhagem receptora com o genótipo *pur*⁻ *pro*⁻ *his*⁻ é infectada e incubada. Depois, os transdutantes para *pur*⁺, *pro*⁺ e *his*⁺ são selecionados individualmente nos experimentos I, II e III, respectivamente.
- a. Que meios são usados para esses experimentos de seleção?
- b. Os transdutantes são examinados quanto à presença de marcadores doadores não-selecionados, com os seguintes resultados:
Qual é ordem dos genes bacterianos?

I		II		III	
<i>pro</i> ⁻ <i>his</i> ⁻	87%	<i>pur</i> ⁻ <i>his</i> ⁻	43%	<i>pur</i> ⁻ <i>pro</i> ⁻	21%
<i>pro</i> ⁺ <i>his</i> ⁻	0%	<i>pur</i> ⁺ <i>his</i> ⁻	0%	<i>pur</i> ⁺ <i>pro</i> ⁻	15%
<i>pro</i> ⁻ <i>his</i> ⁺	10%	<i>pur</i> ⁻ <i>his</i> ⁺	55%	<i>pur</i> ⁻ <i>pro</i> ⁺	60%
<i>pro</i> ⁺ <i>his</i> ⁺	3%	<i>pur</i> ⁺ <i>his</i> ⁺	2%	<i>pur</i> ⁺ <i>pro</i> ⁺	4%

- c. Quais os dois genes mais próximos?
- d. Com base na ordem que você criou na parte c, explique as proporções relativas de genótipos observados no experimento II.
- (O problema 17 é de D. Freifelder, *Molecular Biology and Biochemistry*. Copyright 1978 by W.H. Freeman and Company, New York.)
18. Embora a maioria dos transdutantes *gal*⁺ mediados por λ sejam lisogênicos indutíveis, uma pequena percentagem desses transdutantes de fato não são lisogênicos (isto é, não contêm λ integrada). Os experimentos de controle mostram que esses transdutantes não são produzidos por mutação. Qual a provável origem desses tipos?
19. Uma linhagem bacteriana *ade*⁺ *arg*⁺ *cis*⁺ *his*⁺ *leu*⁺ *pro*⁺ é conhecida como sendo lisogênica para um fago recém-descoberto, mas o sítio do profago não é conhecido. O mapa bacteriano é



A linhagem lisogênica é usada como uma fonte do fago, e os fagos são adicionados em uma linhagem bacteriana do genótipo *ade⁻ arg⁻ cis⁻ his⁻ leu⁻ pro⁻*. Após uma curta incubação, amostras dessas bactérias são plaqueadas em seis meios diferentes, com as suplementações indicadas no quadro seguinte. O quadro também mostra se as colônias foram observadas em vários meios.

Meio	Suplementação de nutriente ao meio						Presença de colônias
	Ade	Arg	Cis	His	Leu	Pro	
1	-	+	+	+	+	+	N
2	+	-	+	+	+	+	N
3	+	+	-	+	+	+	C
4	+	+	+	-	+	+	N
5	+	+	+	+	-	+	C
6	+	+	+	+	+	-	N

(Neste quadro, um sinal + indica a presença de um suplemento nutriente, um sinal - indica que o suplemento não está presente, N indica sem colônias e C indica colônias presentes.)

- a. Que processos genéticos funcionam aqui?
b. Qual o locus aproximado do profago?

20. Em um sistema de transdução generalizada usando o fago P1, o doador é *pur⁺ nad⁺ pdx⁻* e o receptor é *pur⁻ nad⁻ pdx⁺*. O alelo doador *pur⁺* é inicialmente selecionado após a transdução, e 50 transdutantes *pur⁺* são então registrados quanto aos outros alelos presentes. Aqui estão os resultados:

Genótipo	Número de colônias
<i>nad⁺ pdx⁺</i>	3
<i>nad⁺ pdx⁻</i>	10
<i>nad⁻ pdx⁺</i>	24
<i>nad⁻ pdx⁻</i>	13
	50

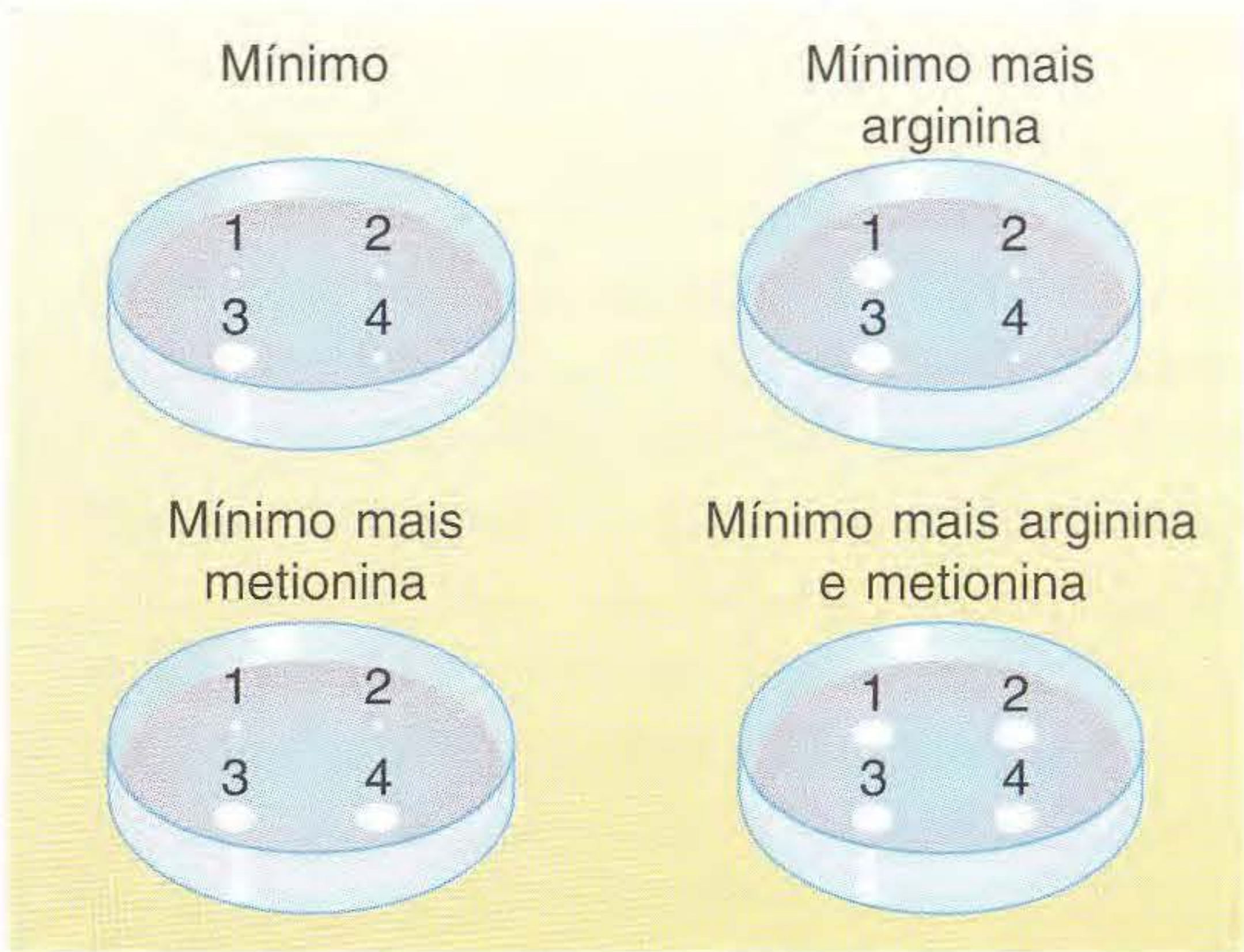
- a. Qual a frequência de co-transdução para *pur* e *nad*?
b. Qual a frequência de co-transdução para *pur* e *pdx*?
c. Qual dos loci não-selecionados é mais próximo de *pur*?
d. *nad* e *pdx* estão do mesmo lado ou em lados opostos a *pur*? Explique. (Desenhe as trocas necessárias para produzir as várias classes transformantes em sua ordem para ver quais precisam do número mínimo para produzir os resultados obtidos.)

21. Em um experimento de transdução generalizada, são coletados fagos de uma linhagem doadora de *E. coli* com o genótipo *cis⁺ leu⁺ tre⁺* e usados para transduzir um genótipo recipiente do genótipo *cis⁻ leu⁻ tre⁻*. Inicialmente, a população receptora tratada é plaqueada em um meio mínimo suplementado com leucina e treonina. São obtidas muitas colônias.

a. Quais os possíveis genótipos dessas colônias?
b. Essas colônias são então submetidas a plaqueamento réplica em três meios diferentes: (1) mínimo mais apenas treonina, (2) mínimo mais apenas leucina e (3) mínimo. Que fenótipos podem, teoricamente, crescer nesses três meios?

- c. Das colônias originais, 56% são observadas crescendo em meio 1, 5% em meio 2 e sem colônias no meio 3. Quais os genótipos reais das colônias nos meios 1, 2 e 3?
d. Desenhe um mapa mostrando a ordem dos três genes e qual dos outros dois genes externos está mais próximo do gene médio.

22. Deduza os genótipos das seguintes linhagens de *E. coli* de 1 a 4:



23. Em um experimento de conjugação ininterrupta em *E. coli*, o gene *pro* entra após o gene *tia*. Uma *pro⁺ tia⁺ Hfr* é cruzada com uma linhagem *pro⁻ tia⁻ F⁻*, e os exconjugantes são plaqueados em um meio contendo tiamina mas não prolina. São observadas 360 colônias, e elas são isoladas e cultivadas em meio totalmente suplementado. Essas culturas são então testadas quanto à sua habilidade em crescer em um meio não contendo prolina ou tiamina (meio mínimo), e 320 das culturas são encontradas sendo capazes de crescer, mas o restante não.
- a. Deduza os genótipos dos dois tipos de culturas.
b. Desenhe os eventos de crossing necessários para produzir esses genótipos.
c. Calcule a distância entre os genes *pro* e *tia* em unidades de recombinação.



Solucionando o Problema 23

- Que tipo de organismo é *E. coli*?
- Com o que se parece uma cultura de *E. coli*?
- Em que tipos de substância geralmente as *E. coli* crescem em seu habitat natural?
- Quais são os requisitos mínimos para a divisão das células *E. coli* ?
- Defina os termos *prototrófico* e *auxotrófico*.
- Que culturas nesse experimento são prototróficas e quais são auxotróficas?
- Considerando algumas linhagens de genótipo desconhecido quanto a tiamina e prolina, como você testaria seus genótipos? Cite os exatos detalhes experimentais, incluindo o equipamento.
- Que tipos de substâncias são prolina e tiamina? Isso é importante nesse experimento?
- Desenhe um diagrama mostrando todas as manipulações feitas no experimento.

10. Por que você acha que o experimento foi feito?
11. Como foi estabelecido que *pro* entra após *tia*? Cite com precisão as etapas experimentais.
12. De que modo um experimento de reprodução interrompida difere do experimento descrito neste problema?
13. O que é um exconjugante? Como você acha que os exconjugantes foram obtidos? (Ele pode incluir genes não descritos neste problema.)
14. Quando o gene *pro* é dito entrar após *tia*, isso significa o alelo *pro*, o alelo *pro*⁺, um deles ou ambos?
15. O que é um “meio totalmente suplementado” no contexto desta questão?
16. Alguns exconjugantes não crescem em meio mínimo. Em que meio eles crescem?
17. Diga os tipos de crossing que tomam parte na recombinação Hfr × F⁻. Como esses crossings diferem dos crossings em eucariontes?
18. O que é uma unidade de recombinação no contexto da presente análise? Como ela difere das unidades de mapa usadas na genética de eucariontes?

24. Um experimento de transdução generalizada usa uma linhagem *metE*⁺ *pirD*⁺ como doadora e *metE*⁻ *pirD*⁻ como receptora. As transdutantes *MetE*⁺ são selecionadas e, então, testadas quanto ao alelo *pirD*⁺. São obtidos os seguintes números:

<i>metE</i> ⁺ <i>pirD</i> ⁻	857
<i>metE</i> ⁺ <i>pirD</i> ⁺	1

Esses resultados sugerem que esses loci são proximalmente ligados? Que outras explicações existem para o “duplo”?

25. Uma linhagem *argC*⁻ foi infectada com um fago transdutor, e o lisado foi usado para transduzir receptores *metF*⁻ em um meio contendo arginina, mas não metionina. As transdutantes *metF*⁺ foram então testadas quanto à necessidade de arginina: a maioria era *argC*⁺, mas uma pequena percentagem era de *argC*⁻. Desenhe diagramas para mostrar a provável origem das linhagens *argC*⁺ e *argC*⁻.

PROBLEMAS DESAFIADORES

26. Quatro linhagens de *E. coli* com o genótipo *a*⁺ *b*⁻ são marcadas 1, 2, 3 e 4. Quatro linhagens com o genótipo *a*⁻ *b*⁺ são marcadas 5, 6, 7 e 8. Os dois genótipos são misturados em todas as combinações possíveis e (após a incubação) são plaqueados para determinar a frequência de recombinantes *a*⁺ *b*⁺. Foram obtidos os seguintes resultados, onde M = muitos recombinantes, L = baixo número de recombinantes e 0 = sem recombinantes:

	1	2	3	4
5	0	M	M	0
6	0	M	M	0
7	L	0	0	M
8	0	L	L	0

Com base nesses resultados, atribua um tipo de sexo (seja Hfr, F⁺ ou F⁻) a cada linhagem.

27. Uma linhagem Hfr com o genótipo *a*⁺ *b*⁺ *c*⁺ *d*⁻ *str*^s é cruzada com uma linhagem fêmea com o genótipo *a*⁻ *b*⁻ *c*⁻ *d*⁺ *str*^r. Em vários momentos, a cultura é agitada vigorosamente para separar os pares reprodutivos. As células foram então plaqueadas em ágar dos seguintes três tipos, onde o nutriente A permite o crescimento de células *a*⁻, o nutriente B, de células *b*⁻, o nutriente C de células *c*⁻, e o nutriente D, de células *d*⁻ (um sinal + indica a presença de estreptomicina ou um nutriente, e um sinal - indica sua ausência):

Tipo de ágar	Str	A	B	C	D
1	+	+	+	-	+
2	+	-	+	+	+
3	+	+	-	+	+

a. Que genes doadores estão sendo selecionados em cada tipo de ágar?

b. O quadro seguinte mostra o número de colônias em cada tipo de ágar para amostras colhidas em vários momentos após as linhagens serem misturadas. Use essa informação para determinar a ordem dos genes *a*, *b* e *c*.

c. De cada uma das placas de 25 min, foram captadas 100 colônias e transferidas para uma placa de Petri contendo ágar com todos os nutrientes, exceto D. Os números de colônias que crescem nesse meio são 89 para a amostra de ágar tipo 1, 51 para a amostra de ágar tipo 2 e 8 para a amostra de ágar tipo 3. Usando esses dados, coloque *d* na sequência de *a*, *b* e *c*.

d. Em que tempo de amostras você espera que as colônias apareçam primeiro em ágar contendo C e estreptomicina, mas não A ou B?

Tempo da amostragem	Número de colônias no ágar tipo		
	1	2	3
0	0	0	0
5	0	0	0
7,5	100	0	0
10	200	0	0
12,5	300	0	75
15	400	0	150
17,5	400	50	225
20	400	100	250
25	400	100	250

(O problema 27 é de D. Freifelder, *Molecular Biology and Biochemistry*. Copyright 1978 by W.H. Freeman and Company.)

28. No cruzamento

$$\text{Hfr } aro^{+} \text{ arg}^{+} \text{ ery}^{r} \text{ str}^{s} \times \text{F}^{-} \text{ } aro^{-} \text{ arg}^{-} \text{ ery}^{s} \text{ str}^{r},$$

os marcadores são transferidos na ordem dada (com *aro*⁺ entrando primeiro), mas os primeiros três genes estão bem próximos. Os exconjugantes são plaqueados em um meio contendo Str (estreptomicina, para matar as células Hfr), Ery (eritromicina), Arg (arginina) e Aro (aminoácidos).

dos aromáticos). Foram obtidos os seguintes resultados de 300 colônias isoladas dessas placas e testadas quanto ao crescimento em vários meios: só em Ery, cresceram 263 linhagens; em Ery + Arg, cresceram 264 linhagens; em Ery + Aro, cresceram 290 linhagens; em Ery + Arg + Aro, cresceram 300 linhagens.

- a. Desenhe uma lista de genótipos e indique o número de indivíduos em cada genótipo.
- b. Calcule as frequências de recombinação.
- c. Calcule a proporção do tamanho da região *arg* para *aro* em relação à região *ery* para *arg*.

29. É feito um experimento de transformação com uma linhagem doadora que é resistente a quatro drogas: A, B, C e D. O receptor é sensível a todas as quatro drogas. A população receptora tratada é dividida e plaqueada em meio contendo várias combinações de drogas. O quadro seguinte mostra os resultados.

Drogas adicionadas	Número de colônias	Drogas adicionadas	Número de colônias
Nenhuma	10.000	BC	51
A	1.156	BD	49
B	1.148	CD	786
C	1.161	ABC	30
D	1.139	ABD	42
AB	46	ACD	630
AC	640	BCD	36
AD	942	ABCD	30

- a. Um dos genes é obviamente bem distante dos outros três, que parecem ser fortemente (proximamente) ligados. Qual o gene mais distante?
 - b. Qual a provável ordem dos três genes proximamente ligados?
- (O problema 29 é de Franklin Stahl, *The Mechanics of Inheritance*, 2nd ed. Copyright 1969, Prentice Hall, Englewood Cliffs, N.J. Reimpresso com permissão.)

30. Você tem duas linhagens de λ que podem lisogenizar *E. coli*; seus mapas de ligação são os seguintes:



O segmento mostrado abaixo do cromossomo, designado 1-2-3, é a região responsável pelo pareamento e crossing com o cromossomo de *E. coli*. (Mantenha os marcadores em todos os seus desenhos.)

- a. Faça um diagrama do modo pelo qual λ linhagem X é inserido no cromossomo de *E. coli* (de modo que *E. coli* é lisogenizada).
- b. As bactérias que são lisogênicas para a linhagem X podem ser superinfectadas usando a linhagem Y. Certa

percentagem dessas bactérias superinfectadas tornam-se “duplamente” lisogênicas (isso é, lisogênicas para ambas as linhagens). Diagrame como isso ocorrerá. (Não se preocupe sobre como as duplo-lisogênicas são detectadas.)

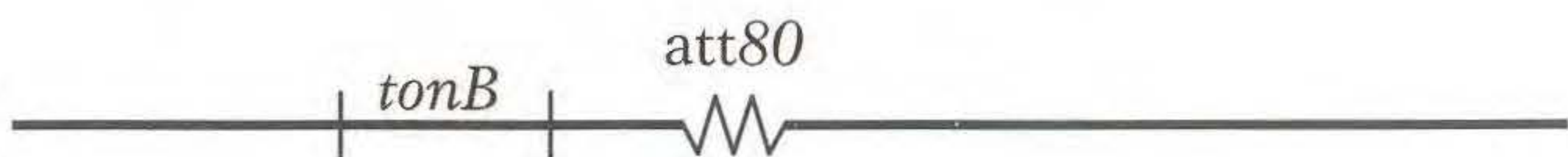
- c. Diagrame como dois profagos λ podem parrear.
- d. Os produtos de crossing entre os dois profagos podem ser recuperados. Diagrame um crossing e as conseqüências.

31. Você tem três linhagens de *E. coli*. A linhagem A é F' *cys*⁺ *trp1*/*cys*⁺ *trp1* (isto é, tanto F' quanto o cromossomo levam *cys*⁺ e *trp1*, um alelo para a necessidade de triptofano). A linhagem B é F⁻ *cys*⁻ *trp2* Z (essa linhagem requer cisteína para crescimento e leva *trp2*, outro alelo causando uma necessidade de triptofano; a linhagem B é lisogênica para o fago Z generalizadamente transduzido). A linhagem C é F⁻ *cys*⁺ *trp1* (é um derivado F⁻ da linhagem A que perdeu o F'). Como você pode determinar se *trp1* e *trp2* são alelos do mesmo locus? (Descreva os cruzamentos e os resultados esperados.)

32. O fago transdutor generalizado é usado para transduzir uma linhagem recipiente de *E. coli* *a*⁻ *b*⁻ *c*⁻ *d*⁻ *e*⁻ com uma doadora *a*⁺ *b*⁺ *c*⁺ *d*⁺ *e*⁺. A cultura receptora é plaqueada em vários meios com os resultados mostrados no quadro seguinte. (Note que *a*⁻ indica necessidade de A como nutriente, e assim por diante.) O que você pode concluir sobre a ligação e a ordem dos genes?

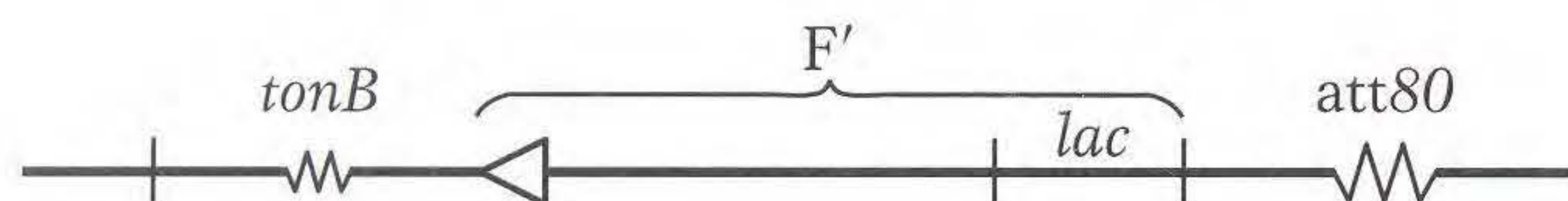
Compostos adicionados ao meio mínimo	Presença (+) ou ausência (-) de colônias
C D E	-
B D E	-
B C E	+
B C D	+
A D E	-
A C E	-
A C D	-
A B E	-
A B D	+
A B C	-

33. Em 1965, Jon Beckwith e Ethan Signer criaram um método para obter fagos transdutores especializados levando a região *lac*. Eles sabiam que o sítio de integração, designado *att80*, para o fago temperado ϕ 80 (um parente do fago λ) estava situado perto de *tonB*, um gene que confere resistência ao fago virulento T1:



Eles usaram o plasmídeo F' *lac*⁺ que não podia replicar-se a altas temperaturas em uma linhagem levando uma deleção dos genes *lac*. Forçando a célula a permanecer *lac*⁺ em altas temperaturas, os pesquisadores puderam selecionar linhagens nas quais o plasmídeo tinha se integrado ao cromossomo, permitindo assim que F' *lac* seja mantido em altas temperaturas. Combinando essa seleção com uma seleção simultânea para resistência à

infecção com o fago T1, eles encontraram que as únicas sobreviventes eram células nas quais a $F' lac$ tinha se integrado ao locus $tonB$, como mostrado aqui:



Esse resultado coloca a região lac perto do sítio de integração para o fago $\phi 80$. Descreva as etapas subsequentes que os pesquisadores devem ter seguido para isolar as partículas transdutoras especializadas do fago $\phi 80$ que leva à região lac .

34. *E. coli* tipo selvagem capta e concentra certo alimento corante em vermelho, tornando as colônias verme-

lho-sangue. Foi usada a mutagênese de transposição, e as células foram plaqueadas em alimento corante. A maioria das colônias eram vermelhas, mas algumas colônias não captaram o corante e se apresentavam brancas. Em uma colônia branca, o DNA que circunda o transposon inserido foi seqüenciado, com o uso de um primer de replicação idêntico à parte do final da seqüência do transposon, e a seqüência adjacente ao transposon era correspondente a um gene de função desconhecida chamado $atoE$, cobrindo as posições 2.322 até 2.324 Mb no mapa (numeradas de uma posição zero arbitrária). Proponha uma função para a $atoE$. Que processo biológico poderia ser investigado desse modo e que outros tipos de colônias brancas podem ser esperadas?

Interação Gênica

6



As cores dos pimentões são determinadas pela interação de vários genes. Um alelo *Y* promove a eliminação inicial da clorofila (um pigmento verde), enquanto *y* não. O alelo *R* determina vermelho e *r* determina o pigmento carotenóide amarelo. Os alelos *c1* e *c2* de dois genes diferentes diminuem as quantidades de carotenóides, causando os tons mais claros. Laranja é a diminuição de vermelho. Marrom é verde mais vermelho. Amarelo-claro é a diminuição de amarelo. [Anthony Griffiths.]

Objeto de nossa apresentação neste livro, até agora, tem sido mostrar como os geneticistas identificam um gene que afeta alguma propriedade biológica de interesse. Vimos como a genética avançada pode ser usada para identificar genes individuais. O pesquisador começa com um conjunto de mutantes, então cruza cada mutante tipo selvagem para ver se o mutante apresenta herança monogênica. Os dados cumulativos de tal programa de pesquisa revelariam um conjunto de genes, todos tendo papéis no desenvolvimento da propriedade em investigação. Em alguns casos, o pesquisador pode ser capaz de identificar funções bioquímicas específicas para muitos dos genes comparando as seqüências gênicas com as de outros organismos. A etapa seguinte, que é um grande desafio, é deduzir como os genes interagem em um conjunto.

A nível molecular, os organismos operam mais ou menos do mesmo modo que uma linha de montagem em uma fábrica, como uma série de vias químicas nas quais cada etapa é controlada por um produto gênico. Existem milhares de tais vias, de tamanhos diferentes e conectadas de vários modos em redes ou siste-

Perguntas Fundamentais

- Como interagem os alelos de um único gene?
- Que são vias multigênicas?
- Como genes separados interagem nessas vias?
- Qual é o teste genético para interação gênica?

Tópicos

- 6.1 Interações de alelos de um único gene: variações de dominância
- 6.2 Interação de genes em vias
- 6.3 Interações gênicas interferentes
- 6.4 Penetrância e expressividade

mas. Os genes que interagem são membros da mesma via ou de uma via conectada. Para fins de discussão, podemos distinguir vários tipos importantes de vias, mas essas vias se superpõem em vários graus. Primeiro, existem *vias puramente biossintéticas*, nas quais a molécula essencial é produzida em uma série de interconversões catalisadas por enzimas codificadas por genes. Segundo, existem *vias de transdução de sinal* que transmitem instruções de um sinal extracelular, como substância ambiental ou um hormônio de outra parte do corpo. Esse sinal ativa uma proteína celular que, por sua vez, ativa outra, e assim por diante, em uma cascata de ativação, a qual geralmente termina na ativação de uma proteína reguladora que liga ou desliga a transcrição de um conjunto de genes. Terceiro, existem *vias desenvolvimentais*, um conjunto diverso de processos, muitos dos quais são controlados por genes, que promovem o crescimento e diferenciação do corpo de um organismo. Quando nós deduzimos tais vias por meio de pesquisa, chegamos perto de compreender o subgrupo do genoma que controla a propriedade que nos interessa.

Consideremos a cor de cabelo humano para ilustrar alguns dos modos pelos quais os genes interagem. O pigmento da cor de cabelo e pele humana é uma substância chamada melanina, que pode ser de dois tipos: eumelanina (tons de preto e castanho) e feomelanina (tons de vermelho e amarelo). A química complicada da melanina é executada nas células da pele chamadas melanócitos. A melanina é sintetizada por uma via bioquímica de várias etapas, sendo cada etapa catalisada por uma enzima. (Lembre que uma das enzimas na via é a tirosinase, codificada pelo gene que, quando defeituoso, produz albinismo, a falta de melanina.) A feomelanina é o produto normal da via biossintética. Ela produz os tons claros dos cabelos encontrados em muitas pessoas no mundo. Entretanto, a maior parte da população mundial tem cabelos escuros. Nessas pessoas, a feomelanina é convertida em eumelanina. Essa conversão requer a ação de um hormônio proteico, o hormônio estimulante do melanócito (MSH), codificado por um gene transcrito na hipófise. O MSH surge em um melanócito e liga-se ao receptor de melanocortina-1. A ligação de MSH a seu receptor ativa um processo de tradução de sinal no melanócito que resulta na produção de eumelanina. Assim, nesse estudo de caso descrito, vemos uma interação de três tipos de vias. Os melanócitos devem ser diferenciados por uma *via desenvolvimental*, a melanina é produzida por uma *via sintética* e a conversão da feomelanina requer uma *transdução de sinal* iniciada por hormônio. Todos esses processos são controlados por genes.

Adicional a essa complexidade está a sutileza baseada nos alelos de um único gene. Vários alelos do gene do receptor de melanocortina-1 (*mc1r*) estão presentes nas populações humanas, e apenas um dos alelos resulta eumelanina, necessária para o cabelo escuro. Os outros alelos resultam em várias formas de feomelanina. Eles resultam em cabelo ruivo (Figura 6.1), cabelo castanho-claro e cabelo loiro. Essas interações de genes determinando a cor do cabelo humano foram deduzidas, em grande parte, de estudos da interação gênica em camundongos, o principal modelo genético para mamíferos, em geral, e humanos, especificamente.

Como são deduzidas as interações subjacentes a uma propriedade? Um enfoque é analisar diretamente as interações

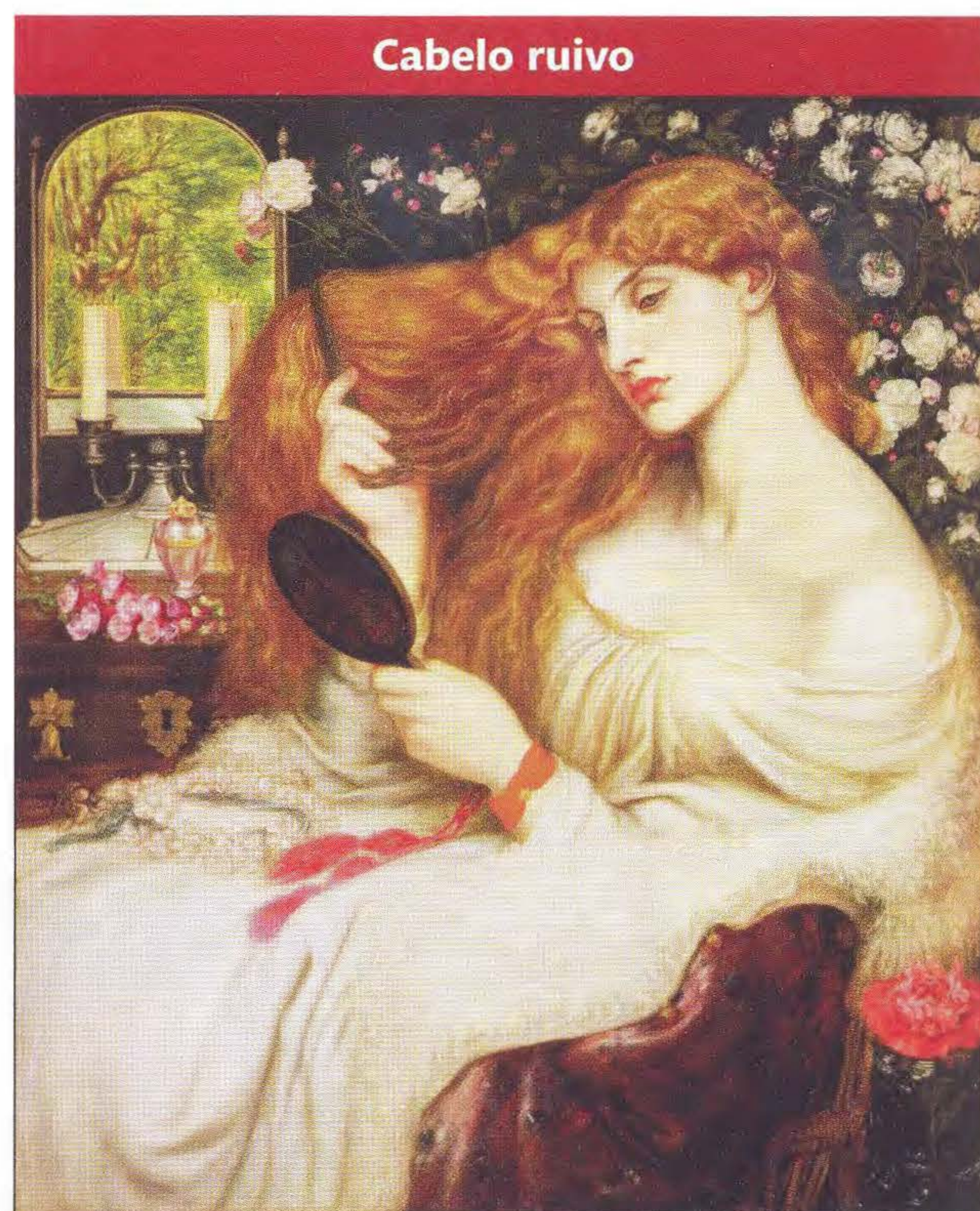


FIG. 6.1 A expressão do cabelo ruivo, herdado como autossômico recessivo, em uma análise mais aprofundada requer a interação de muitos outros genes. [Copyright Delaware Art Museum, Wilmington, USA/Samuel and Mary R. Bancroft Memorial /The Bridgeman Art Library.]

ções de proteínas *in vitro* usando uma proteína como “isca” e observando que outras proteínas celulares ligam-se a ela. As proteínas que são encontradas ligando-se podem ser componentes de uma “máquina” celular proteica. Outro enfoque é analisar os transcritos de mRNA. Os genes que colaboram em algum processo desenvolvimental específico podem ser definidos pelo conjunto de RNA transcritos presentes quando esse processo está ocorrendo, um tipo de análise hoje feita com uso de chips genômicos (veja Capítulo 13). Finalmente, a interação gênica pode ser deduzida pela *análise genética* direta, que é o foco deste capítulo.

Como funciona um enfoque de genética direta na interação? Traçaremos a seqüência comum de etapas em um programa de pesquisa genética. Primeiro, como descrito nos capítulos anteriores, os mutantes são cruzados com o tipo selvagem para ver se os mutantes apresentam herança monogênica. Os padrões de herança devem revelar, simultaneamente, se a nova mutação é recessiva ou dominante. Segundo, a possibilidade de múltiplos eventos no mesmo gene deve ser resolvida. A pergunta principal seria: Qualquer mutação no conjunto seria de mutações no mesmo gene? Terceiro, as mutações são testadas aos pares para ver se elas interagem, fazendo *mutantes duplos* dos genes envolvidos. A idéia é que, se os genes interagem, então o mutante duplo mostrará um fenótipo que difere dos fenótipos mutantes individuais, aparecendo juntos no mesmo organismo individual. Os mutantes duplos são recuperados da geração F_2 . Na F_2 , a proporção de fenótipos diferentes pode ajudar a iden-

tificar o mutante duplo em avaliação. Além disso, nos casos nos quais os dois alelos mutantes interagem, resultará uma proporção mendeliana modificada de 9:3:3:1.

6.1 Interações de Alelos de um Único Gene: Variações de Dominância

Começamos com uma discussão geral da natureza e ação dos alelos. Existem milhares de modos diferentes de alterar a sequência de um gene, produzindo uma mutação, embora apenas alguns desses alelos mutantes apareçam em uma população real. Os mutantes conhecidos de um gene são chamados de **alelos múltiplos** ou uma **série alélica**.

Um dos testes feitos em um alelo mutante é ver se ele é dominante ou recessivo. A informação básica sobre dominância e recessividade é útil para trabalhar com a mutação nova, e pode ser uma fonte de esclarecimento de como a mutação atua, como veremos nos exemplos. A dominância é realmente um tipo de interação *de alelos de um único gene* em um heterozigoto. Os alelos que interagem podem ser do tipo selvagem e alelos mutantes (+/m) ou dois alelos mutantes diferentes (m_1/m_2). De fato, existem vários tipos de dominância, cada uma representando um tipo diferente de interação de alelos.

Dominância completa e recessividade

O tipo mais simples de dominância é a **dominância total** ou **completa**. Um alelo completamente dominante será expres-

so quando apenas uma cópia está presente, como no heterozigoto, enquanto o alelo alternativo será totalmente recessivo. Na dominância completa, o homozigoto dominante não pode ser distinto do heterozigoto; isto é, a nível fenotípico, $A/A = A/a$. Como mencionado antes, a fenilcetonúria (PKU) e muitas outras doenças humanas monogênicas são totalmente recessivas, enquanto seus alelos tipo selvagem são dominantes. Outras doenças monogênicas, tais como acondroplasia, são totalmente dominantes, embora, nesses casos, o alelo tipo selvagem seja recessivo. Como podem essas relações de dominância ser interpretadas a nível celular?

A doença PKU é um bom modelo geral para mutações recessivas. Lembre que a PKU é causada por alelo defeituoso do gene codificante da enzima fenilalanina hidroxilase (PAH). Na ausência de PAH normal, a fenilalanina que entra no corpo pelos alimentos não é degradada, e, portanto, acumula-se. Sob tais condições, a fenilalanina é convertida em ácido pirúvico, que é transportado para o cérebro pela corrente sanguínea, onde impede o desenvolvimento normal, levando ao retardo mental. O motivo de o alelo defeituoso ser recessivo é que uma “dose” do alelo tipo selvagem *P* produz PAH suficiente para degradar a fenilalanina que entra no corpo. O gene PAH é dito *haplossuficiente*. Assim, tanto P/P (duas doses) quanto P/p (uma dose) têm atividade suficiente de PAH para resultar na química celular normal. As pessoas com p/p têm uma dose zero de atividade de PAH. A Figura 6.2 ilustra essa noção geral.

Mutações de genes haplossuficientes são recessivas

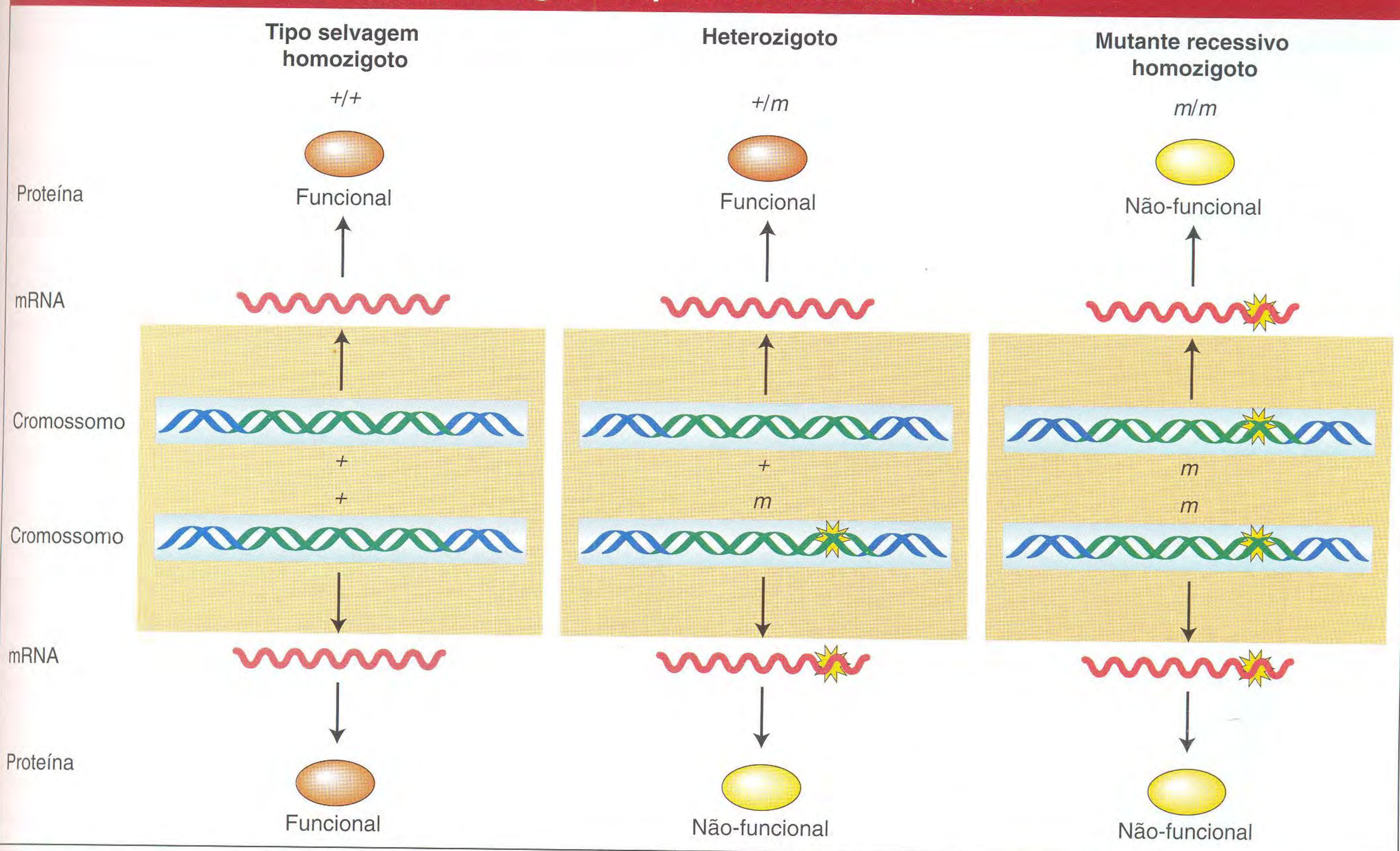


FIG. 6.2 No heterozigoto, muito embora a cópia mutada do gene produza proteína não-funcional, a cópia tipo selvagem gera suficiente proteína funcional para produzir o fenótipo tipo selvagem.

Como podemos explicar mutações totalmente dominantes? Existem vários mecanismos moleculares para dominância. Um mecanismo comumente encontrado é que o alelo tipo selvagem de um gene é *haploinsuficiente*. Na haploinsuficiência, uma dose do tipo selvagem *não* é suficiente para atingir os níveis normais de função. Suponha que 16 unidades do produto de um gene sejam necessárias para a química normal e que cada alelo tipo selvagem possa fazer 10 unidades. Dois alelos tipo selvagem produzem 20 unidades do produto, bem acima do mínimo. Mas considere o que ocorre se uma das mutações é uma **mutação nula**, que produz uma proteína não-funcional. Uma mutação nula em combinação com um único alelo tipo selvagem produziria $10 + 0 = 10$ unidades, bem abaixo do mínimo. Assim, o heterozigoto (tipo selvagem/nulo) é mutante, e a mutação é, por definição, dominante. Em camundongos, o gene *Tbx1* é haploinsuficiente. Esse gene codifica uma proteína reguladora de transcrição (um *fator de transcrição*) que atua nos genes responsáveis pelo desenvolvimento da faringe. Um nocaute de um dos alelos tipo selvagem resulta em uma concentração inadequada da proteína reguladora, que resulta em defeitos no desenvolvimento das artérias da faringe. A mesma haploinsuficiência é tida como responsável pela síndrome de DiGeorge em humanos, uma condição com anomalias cardiovasculares e craniofaciais.

Outro tipo importante de mutação dominante é chamada de **dominante negativa**. Os polipeptídeos com esse tipo

de mutação atuam como “*spoilers*” ou “marotas”. Em alguns casos, o produto gênico é uma unidade de uma proteína *homodimérica*, uma proteína composta de duas unidades do mesmo tipo. No heterozigoto (+/*M*), o polipeptídeo *spoiler* liga-se ao polipeptídeo tipo selvagem e o distorce ou interfere em sua função. O mesmo tipo de estrago também pode prejudicar o funcionamento de um composto *heterodímero* de polipeptídeos de genes diferentes. Em outros casos, o produto gênico é um monômero, e, nessas situações, o mutante liga-se ao substrato, deixando pouca disponibilidade na qual a proteína tipo selvagem possa agir.

Um exemplo de mutações que podem agir como dominantes negativas é encontrado no gene para a proteína do colágeno. Algumas mutações nesse gene dão origem ao fenótipo humano osteogênese imperfeita (doença dos ossos quebradiços). O colágeno é uma proteína do tecido conjuntivo formada por três monômeros entrelaçados (um trímero). No heterozigoto mutante, a proteína anormal enrola-se ao redor de uma ou duas normais e distorce o trímero, levando ao mal funcionamento. Desse modo, o colágeno defeituoso atua como um *spoiler* (estragador). A diferença entre a haploinsuficiência e a ação de um dominante negativo como causa da dominância de uma mutação é ilustrada na Figura 6.3.

Para a maioria dos genes, uma única cópia é adequada para a expressão total (tais genes são haplossuficientes), e suas mutações nulas são totalmente recessivas. Mutações prejudiciais de genes haploinsuficientes são, em geral, dominantes. As mutações nos genes que codificam unidades em homou heterodímeros podem se comportar como dominantes negativas, agindo por proteínas “*spoiler*”.

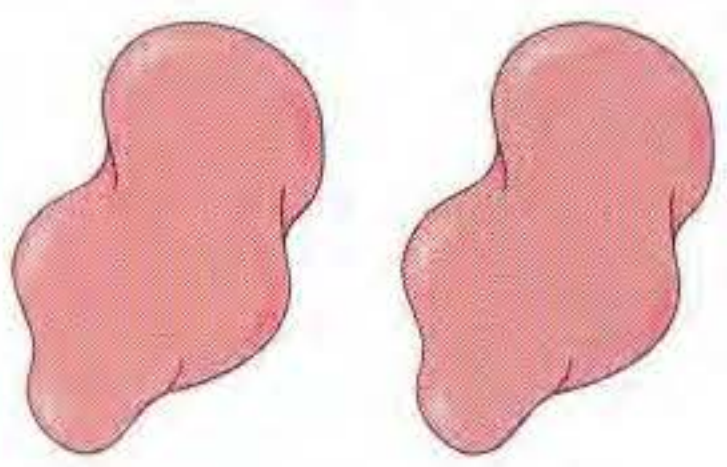
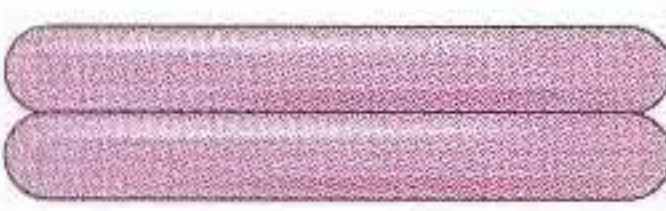

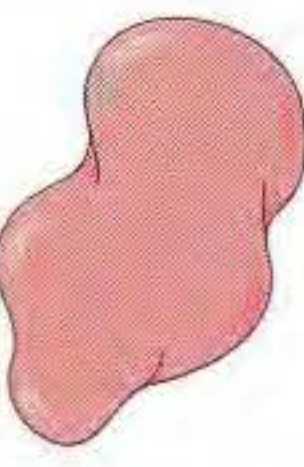
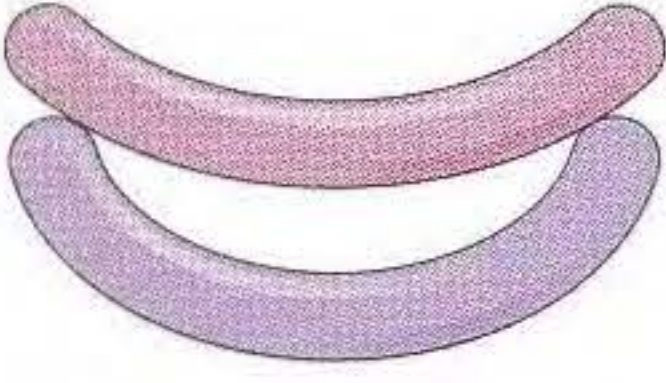
Dois modelos de dominância de uma mutação		
	Haploinsuficiência	Dominante negativo
+/+	 2 “doses” do produto	 Dímero
M/M	 0 “dose”	
+/M	 1 “dose” (inadequada)	

FIG. 6.3 Uma mutação pode ser dominante (*esquerda*) porque um único gene tipo selvagem não produz proteína suficiente para o funcionamento apropriado, ou (*direita*) o alelo mutante atua como um dominante negativo que produz um produto proteico “prejudicial” (*spoiler*).

Dominância incompleta

A maravilha é uma planta nativa da América tropical. Seu nome (*four-o'clock*) vem do fato de que suas flores se abrem ao final da tarde. Quando uma linhagem tipo selvagem pura com pétalas vermelhas é cruzada com uma linhagem pura com pétalas brancas, a F₁ tem pétalas rosa. Se for produzida uma F₂ por autofecundação da F₁, o resultado é

- 1/4 de plantas com pétalas vermelhas
- 1/2 de plantas com pétalas rosa
- 1/4 de plantas com pétalas brancas

A Figura 6.4 mostra esses fenótipos. Por esta proporção 1:2:1 na F₂, podemos deduzir que o padrão de herança é baseado em dois alelos de um único gene. Entretanto, os heterozigotos (a F₁ e metade da F₂) são de fenótipo intermediário. Inventando símbolos alélicos, podemos listar os genótipos da maravilha nesse experimento como *c*⁺/*c*⁺ (vermelha), *c*/*c* (branca) e *c*⁺/*c* (rosa). A ocorrência do fenótipo intermediário sugere uma **dominância incompleta**, o termo usado para descrever o caso geral no qual o fenótipo de um heterozigoto é intermediário ao de dois homozigotos, em alguma escala quantitativa de medida.

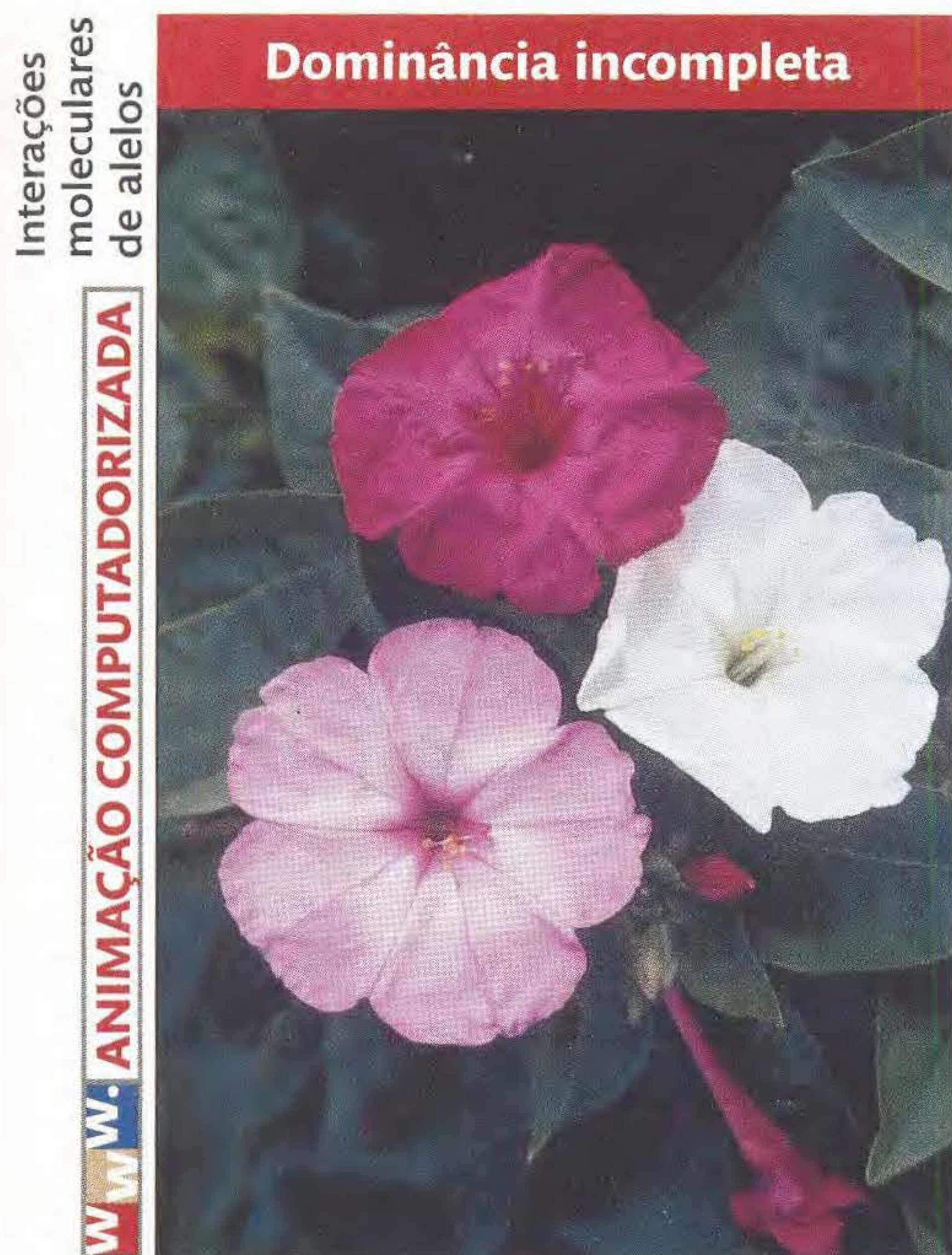


FIG. 6.4 Na planta maravilha, um heterozigoto é rosa, intermediário entre os dois homozigotos vermelho e branco. O heterozigoto rosa demonstra dominância incompleta. [R. Calentine/Visuals Unlimited.]

Como explicamos a dominância incompleta a nível molecular? Na dominância incompleta, cada alelo tipo selvagem produz uma dose de seu produto proteico. O número de doses de um alelo tipo selvagem determina a concentração de uma substância feita pela proteína, tal como um pigmento. Na planta maravilha, duas doses produzem a maioria das cópias do transcrito, gerando assim a maior quantidade de proteína e, portanto, a maior quantidade do pigmento, bastante para fazer as pétalas serem vermelhas. Uma dose produz menos pigmento, e, assim, as pétalas são rosa. Uma dose zero não produz pigmento.

Co-dominância

Outra variação do tema da dominância é a **co-dominância**, a expressão de ambos os alelos de um heterozigoto. Um exemplo claro é visto nos grupos sanguíneos ABO de humanos, onde há co-dominância de alelos de antígeno. Os grupos sanguíneos ABO são determinados por três alelos de um gene. Os três alelos interagem de vários modos para produzir os quatro tipos de sistema sanguíneo ABO. Os três principais alelos são i , I^A e I^B , mas uma pessoa pode ter apenas dois dos três alelos ou duas cópias de um deles. As combinações resultam em seis genótipos diferentes: os três homozigotos e três tipos diferentes de heterozigotos, do seguinte modo:

Genótipo	Tipo sanguíneo
I^A/I^A , I^A/i	A
I^B/I^B , I^B/i	B
I^A/I^B	AB
i/i	O

Nessa série alélica, os alelos determinam a presença e forma de uma molécula de açúcar complexa presente na superfície das hemácias. Essa molécula de açúcar é um antígeno, uma molécula de superfície celular que pode ser reconhecida pelo sistema imunológico. Os alelos I^A e I^B determinam duas formas diferentes de molécula de superfície celular. Entretanto, o alelo i resulta na falta de molécula de superfície desse tipo (é um alelo nulo). Nos genótipos I^A/i e I^B/i , os alelos I^A e I^B são totalmente dominantes em relação a i . Entretanto, no genótipo I^A/I^B , cada um dos alelos produz sua própria forma de molécula de superfície celular, e, assim, os alelos A e B são co-dominantes.

A doença humana anemia falciforme fornece um exemplo interessante das vias um tanto arbitrárias nas quais classificamos a dominância. O gene envolvido codifica uma molécula de globina, que é responsável pelo transporte de oxigênio nos vasos sanguíneos e é o principal constituinte das hemácias. Existem dois alelos principais, Hb^A e Hb^S , e os três genótipos possíveis têm fenótipos diferentes, como em seguida:

- Hb^A/Hb^A : normal; as hemácias nunca afoiçam
- Hb^S/Hb^S : anemia grave, geralmente fatal; a hemoglobina anormal faz com que as hemácias fiquem falciformes
- Hb^A/Hb^S : sem anemia; as hemácias só afoiçam sob baixas concentrações de oxigênio

A Figura 6.5 mostra uma micrografia eletrônica de hemácias incluindo algumas células afoiçadas. Com relação à presença ou ausência de anemia, o alelo Hb^A é dominante. No heterozigoto, um único alelo Hb^A produz suficiente hemoglobina funcional para impedir a anemia. Com relação à forma da hemácia, há uma dominância incompleta, como mostrado pelo fato de que, no heterozigoto, muitas das células têm uma forma um pouco afoiçada. Finalmente, com relação à hemoglobina em si, há uma co-dominância. Os alelos Hb^A e Hb^S codificam duas formas diferentes de hemoglobina que diferem por um único aminoácido, e ambas as formas são

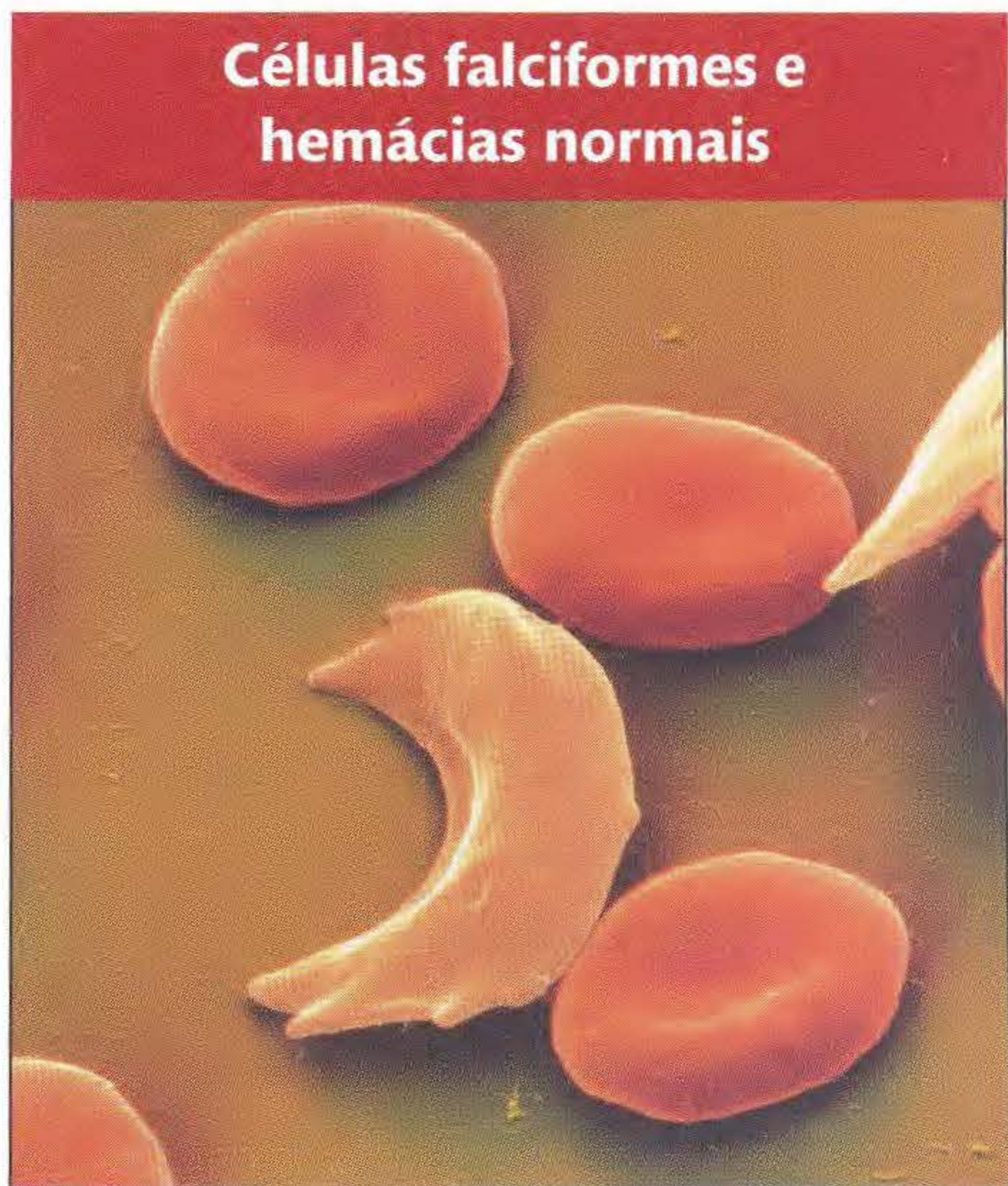


FIG. 6.5 A célula falciforme é causada por uma mutação no gene para hemoglobina. [Meckes/Ottawa/Photo Researchers.]

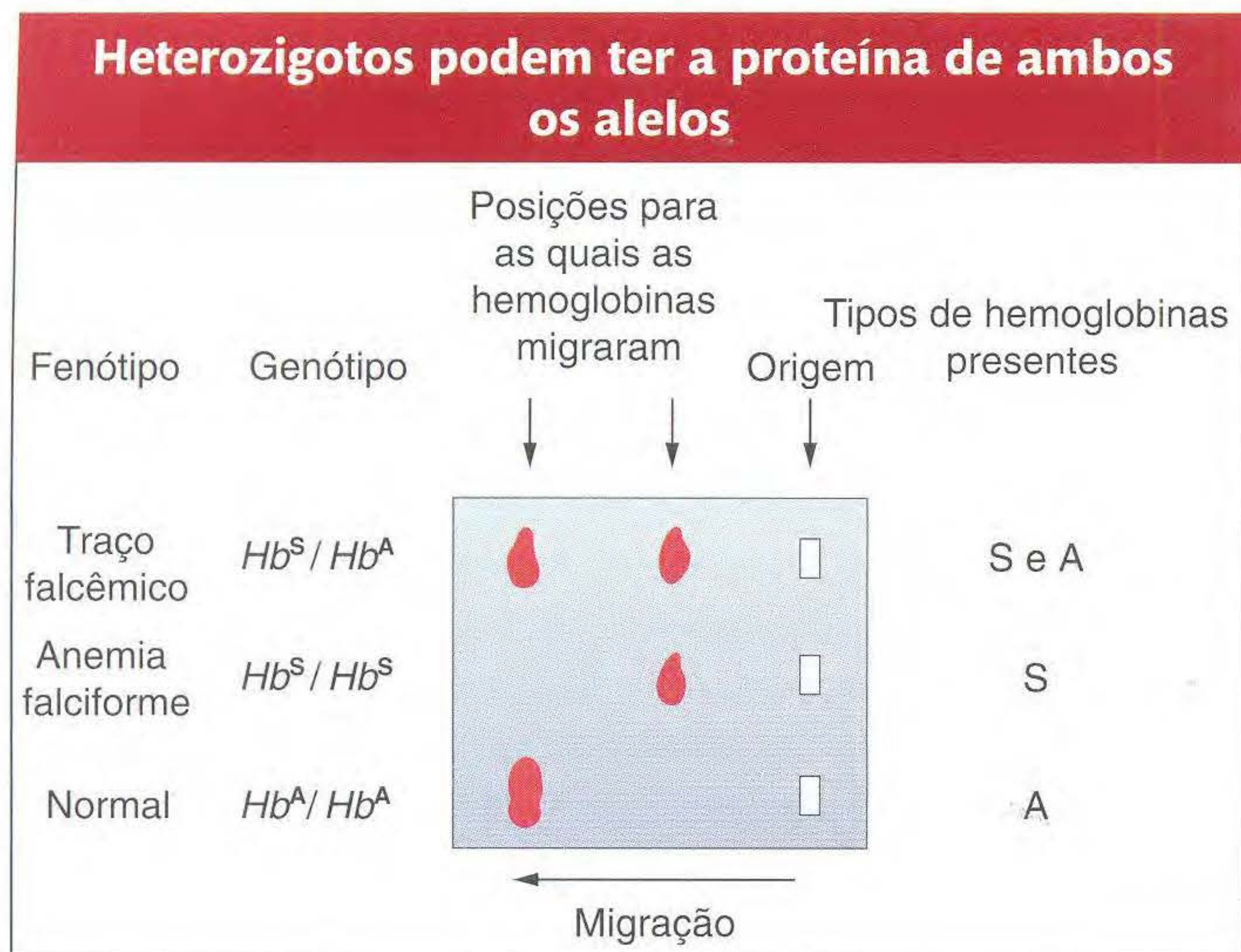


FIG. 6.6 A eletroforese das hemoglobinas normal e mutante. São mostrados os resultados produzidos pela hemoglobina de uma pessoa com traço falcêmico (um heterozigoto), uma pessoa com anemia falciforme e uma pessoa normal. As manchas mostram as posições para as quais as hemoglobinas migraram no gel de amido.

produzidas pelo heterozigoto. As formas A e S de hemoglobina podem ser separadas por eletroforese, pois ocorre que elas têm cargas diferentes (Figura 6.6). Vemos que as pessoas normais homozigotas têm um tipo de hemoglobina (A) e os anêmicos têm outro (tipo S), que se move mais lentamente no campo elétrico. Os heterozigotos têm ambos os tipos, A e S. Em outras palavras, há co-dominância a nível molecular. A fascinante genética de populações dos alelos Hb^A e Hb^S será considerada no Capítulo 16.

A anemia falciforme ilustra a arbitrariedade dos termos *dominância*, *dominância incompleta* e *co-dominância*. O tipo de dominância deduzida depende do nível fenotípico no qual a análise é feita — do organismo, da célula ou da molécula. De fato, deve ser aplicada certa cautela a muitas das categorias que os cientistas usam para classificar as estruturas e processos; essas categorias são criadas pelos humanos para conveniência da análise.

O tipo de dominância é determinado pelas funções moleculares dos alelos de um gene e pelo nível investigativo da análise.

As folhas da planta trevo apresentam muitas variações quanto ao tema dominância. Trevo é o nome comum para as plantas do gênero *Trifolium*. Existem muitas espécies. Algumas são nativas da América do Norte, enquanto outras aí crescem como pragas introduzidas. Muitas pesquisas genéticas foram feitas com o trevo branco, que apresenta considerável variação entre plantas individuais no curioso padrão em V (divisas) das folhas. As diferentes formas de divisas (e a ausência de divisas) são determinadas por uma série alélica, como visto na Figura 6.7, que mostra os muitos tipos de interações possíveis para um mesmo alelo.

Sete alelos e suas interações no padrão de folhas de trevo

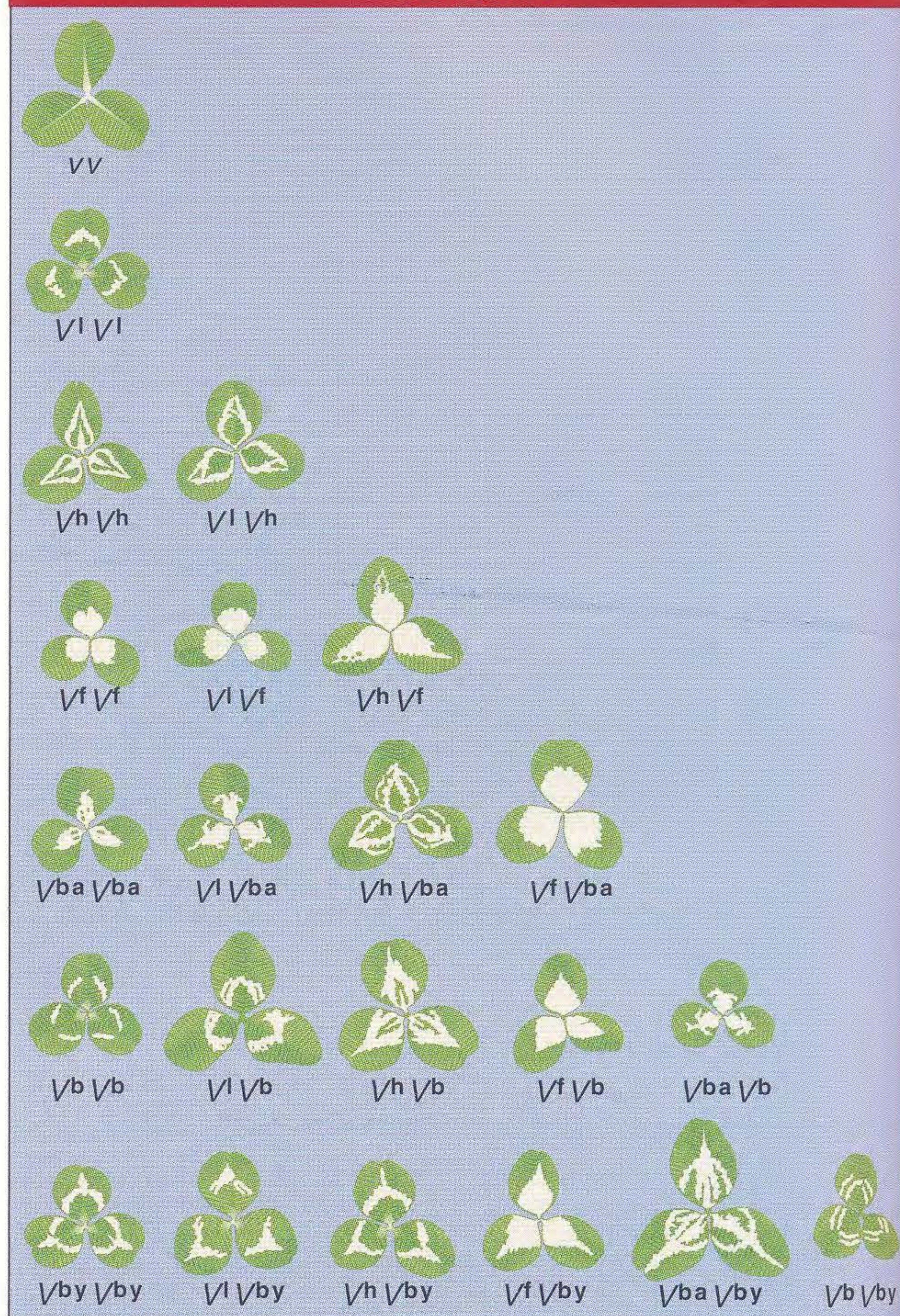


FIG. 6.7 Os alelos múltiplos determinam o padrão (divisas) nas folhas da planta branca. O genótipo de cada planta é mostrado abaixo dela. Há uma variedade de interações de dominância. [Foto de W. Ellis Davies.]

Alelos letais recessivos

Um alelo que é capaz de causar a morte de um organismo é chamado de **alelo letal**. Na caracterização de um conjunto de alelos mutantes recém-descobertos, uma mutação recessiva é às vezes descoberta como sendo letal. Essa informação é potencialmente útil, pois mostra que o gene recém-descoberto (de função ainda desconhecida) é essencial para a operação do organismo. De fato, com o uso da moderna tecnologia do DNA, um alelo mutante nulo de um gene de interesse agora pode ser feito intencionalmente, e tornado homozigoto para ver se ele é letal e sob quais condições ambientais. Os alelos letais também são úteis para determinar o estágio desenvolvimental no qual o gene normalmente atua. Nesse caso, os geneticistas investigam se a morte ocorreu por um alelo mutante letal cedo ou tarde no desenvolvimento de um zigoto. O fenótipo associado à morte também pode ser informativo com relação à função do gene; por exemplo, se certo órgão parece ser anormal, o gene provavelmente se expressa nesse órgão.

Qual o teste diagnóstico para a letalidade? O teste é bem ilustrado por um dos exemplos prototípicos de um alelo letal,

Um alelo letal recessivo, pelagem bege



FIG. 6.8 Uma ninhada de um cruzamento entre dois camundongos heterozigotos para o alelo dominante de cor de pelagem bege. O alelo é letal em dose dupla. Nem toda a prole é visível. [Anthony Griffiths.]

um alelo de cor de pelagem em camundongo (veja o boxe de Organismo-modelo adiante). Os camundongos normais tipo selvagem têm pelagem com uma pigmentação geral escura. Uma mutação chamada de *yellow* (uma cor mais clara) mostra um curioso padrão de herança. Se qualquer camundongo *yellow* é cruzado com um homozigoto tipo selvagem, sempre é observada na prole uma proporção de 1:1 *yellow* para tipo selvagem. Esse resultado sugere que um camundongo *yellow* é sempre heterozigoto para o alelo *yellow*, e o alelo *yellow* é dominante em relação ao tipo selvagem. Entretanto, se quaisquer dois camundongos *yellow* são cruzados entre si, o resultado é sempre o seguinte:

$$\text{yellow} \times \text{yellow} \rightarrow \frac{2}{3} \text{yellow}, \frac{1}{3} \text{tipo selvagem}$$

A Figura 6.8 mostra uma ninhada típica de um cruzamento entre camundongos *yellow*.

Como a proporção 2:1 pode ser explicada? Esses resultados fazem sentido se o alelo *yellow* supostamente é letal quando homozigoto. O alelo *yellow* é conhecido como um gene de cor de pelagem chamado A^Y , que vamos chamar de A^Y . Assim, os resultados do crossing de dois camundongos *yellow* são

	$A^Y/A \times A^Y/A$	
Prole	$\frac{1}{4} A^Y/A^Y$	letal
	$\frac{1}{2} A^Y/A$	<i>yellow</i>
	$\frac{1}{4} A/A$	tipo selvagem

A proporção monoíbrida esperada de 1:2:1 seria encontrada entre os zigotos, mas é alterada para 2:1 na prole vista ao nascimento porque os zigotos com um genótipo letal A^Y/A^Y não sobrevivem para serem contados. Essa hipótese é apoiada pela remoção do útero de fêmeas grávidas do cruzamento *yellow* \times *yellow*; 25% dos embriões são encontrados mortos.

O alelo A^Y produz efeitos em duas características: cor da pelagem e sobrevida. Entretanto, é totalmente possível que ambos os efeitos do alelo A^Y resultem da mesma causa básica, o que causa pelagem *yellow* em uma só dose e morte em dose dupla. Em geral, o termo **pleiotrópico** é usado para qualquer alelo que afete várias propriedades de um organismo.

Sem cauda, um alelo letal recessivo em gatos



FIG. 6.9 Um gato Manx. Um alelo dominante causador de falta de cauda é letal no estado homozigoto. O fenótipo de duas cores de olho não tem relação com a falta de cauda. [Gerard Lacz/NHPA.]

O fenótipo sem cauda Manx em gatos (Figura 6.9) também é produzido por um alelo que é letal no estado homozigoto. Uma única dose do alelo Manx, M^L , interfere gravemente no desenvolvimento da coluna dorsal, resultando na falta de cauda no heterozigoto M^L/M . Mas, no homozigoto M^L/M^L , a dupla dose do gene produz uma anomalia tão extrema no desenvolvimento da coluna que o embrião não sobrevive.

Os alelos *yellow* e M^L têm seus próprios fenótipos em um heterozigoto, mas a maioria dos letais recessivos são silenciosos no heterozigoto. Em tal situação, a letalidade recessiva é diagnosticada observando a morte de 25% da prole em algum estágio do desenvolvimento.

Um alelo ser letal ou não depende do ambiente no qual o organismo se desenvolve. Enquanto alguns alelos são letais em praticamente qualquer ambiente, outros são viáveis em um ambiente mas letais em outro. As doenças hereditárias humanas fornecem alguns exemplos. A fibrose cística e a anemia falciforme são doenças que seriam letais sem tratamento. Além disso, muitos dos alelos favorecidos e selecionados por criadores de animais e plantas quase que certamente seriam eliminados na natureza como resultado de competição com membros da população natural. As variedades mutantes de grão anão, que têm alta produtividade, fornecem bons exemplos; apenas uma cuidadosa nutrição pelos fazendeiros manteve tais alelos para nosso benefício.

Os geneticistas comumente encontram situações nas quais as proporções fenotípicas esperadas são consistentemente desviadas em uma direção porque um alelo mutante reduz a viabilidade. Por exemplo, no cruzamento $A/a \times a/a$, prevemos uma proporção de prole de 50% A/a e 50% a/a , mas podemos consistentemente observar uma proporção tal como 55%:45% ou 60%:40%. Em tal caso, o alelo recessivo é dito como sendo *subletal* porque a letalidade é expressa apenas em alguns indivíduos homozigotos, mas não em todos.



Organismo-modelo Camundongo

O camundongo de laboratório é descendente do camundongo caseiro *Mus musculus*. As linhagens puras usadas hoje como padrões são derivadas de camundongos cruzados nos últimos séculos pelos apreciadores de camundongo. Entre os organismos-modelo, há um cujo genoma mais proximamente se assemelha ao genoma humano. Seu número diplóide de cromossomos é 40 (comparado com os 46 em humanos), e o genoma é um pouco menor que o dos humanos (o genoma humano tem 3.000 Mb) e contém aproximadamente o mesmo número de genes (a estimativa atual é de 25.000). Além disso, todos os cromossomos de camundongos parecem ter contrapartes nos humanos. Uma grande proporção dos genes estão dispostos em blocos exatamente nas mesmas posições que os humanos.

A pesquisa de genética mendeliana de camundongos começou no início do século XX. Uma das mais importantes contribuições iniciais foi a elucidação dos genes que controlam a cor e o padrão da pelagem. O controle genético da pelagem de camundongo forneceu um modelo para todos os mamíferos, incluindo gatos, cães, cavalos, e gado. Também foi dedicada uma grande quantidade de trabalho com mutações induzidas por radiação e substâncias químicas. A genética de camundongos tem tido um grande significado em medicina. Uma grande proporção de doenças genéticas humanas têm contrapartes úteis para estudos experimentais (elas são chamadas de “modelos animais”). O camundongo teve um papel particularmente importante no desenvolvimento de nossa atual compreensão dos genes subjacentes ao câncer.

O genoma do camundongo pode ser modificado pela inserção de fragmentos específicos de DNA em um zigoto ou em células somáticas. O camundongo da foto recebeu um gene de água-viva para uma proteína fluorescente verde (GFP) que o faz brilhar em verde. Os nocautes gênicos e substituições também são possíveis.

Uma grande limitação da genética de camundongos é seu custo. Enquanto trabalhar com um milhão de bactérias

E. coli ou *S. cerevisiae* é uma questão trivial, trabalhar com um milhão de camundongos requer uma construção do tamanho de uma fábrica. Além disso, embora os camundongos se reproduzam com rapidez (comparados com os humanos), eles não podem competir com microrganismos quanto à rapidez do ciclo vital. Assim, as seleções em larga escala e triagens necessárias para detectar eventos raros não são possíveis.



Embriões de camundongo geneticamente modificados fluorescendo em verde. O gene de água-viva para uma proteína fluorescente foi inserido no cromossomo do camundongo brilhante. Os camundongos normais são arroxeados. [Kyodo News.]

Assim, a letalidade pode variar de 0 a 100%, dependendo do próprio gene, do resto do genoma e do ambiente.

Vimos que os genes letais são úteis para diagnosticar a época em que um gene atua e a natureza do defeito fenotípico que mata. Entretanto, a manutenção dos estoques portadores de alelos letais para uso em laboratório é um desafio. Nos diplóides, os alelos letais recessivos podem ser mantidos como heterozigotos. Nos haplóides, os alelos letais sensíveis ao calor são úteis. Eles são membros de uma classe geral de **mutações sensíveis à temperatura (ts)**. Seu fenótipo é tipo selvagem na **temperatura permissiva** (temperatura ambiente em geral), mas mutante em alguma **temperatura restritiva** mais alta. Os alelos sensíveis à temperatura são tidos como

causados por mutações que tornam a proteína propensa a torcer ou dobrar sua forma para uma conformação inativa na temperatura restritiva. Os estoques de pesquisa podem ser mantidos facilmente sob condições permissivas, e o fenótipo mutante pode ser triado em um subgrupo de indivíduos por uma mudança para as condições restritivas. As mutações letais dominantes sensíveis à temperatura também são úteis. O tipo de mutação é expresso mesmo quando presente em uma única dose, mas apenas quando os experimentadores mudam o organismo para a temperatura restritiva.

Os alelos nulos para genes identificados por seqüenciamento genômico podem ser feitos usando uma variedade de procedimentos de “genética reversa” que especificamente

nocauteiam o funcionamento desse gene. Eles serão descritos no Capítulo 13.

Para ver se um gene é essencial, um alelo nulo é testado quanto à letalidade.

6.2 Interação de Genes em Vias

Os genes atuam controlando a química celular. No início do século XX, Archibald Garrod, um médico inglês (Figura 6.10), fez a primeira observação que apoiou esse *insight*. Garrod notou que várias doenças humanas recessivas apresentavam defeitos no que é chamado de metabolismo, o conjunto geral de reações químicas que ocorrem em um organismo. Essa observação levou à noção de que tais doenças genéticas são “erros inatos do metabolismo”. Garrod trabalhou em uma doença chamada de alcaptonúria (AKU), ou doença da urina preta. Ele descobriu que a substância responsável pela urina preta era o ácido homogentísico, que está presente em grandes quantidades e é secretado na urina dos pacientes AKU. Ele sabia que, nas pessoas não-afetadas, o ácido homogentísico é convertido em ácido maleilacetoacético; então, propôs ele que, na AKU, há um defeito nessa conversão. Conseqüentemente, o ácido homogentísico acumula-se e é excretado. As observações de Garrod levantaram a possibilidade de que as vias químicas da célula estavam sob o controle de um grande conjunto de genes interagindo. Entretanto, a demonstração direta desse controle foi dada pelo trabalho posterior de Beadle e Tatum.

Vias biossintéticas em *Neurospora*

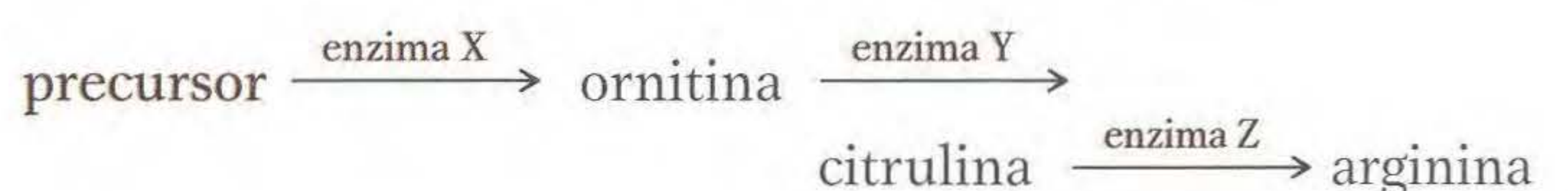
O estudo de George Beadle e Edward Tatum, que foi um marco nos anos 1940, não só esclareceu o papel dos genes, mas tam-

bém demonstrou a interação dos genes nas vias bioquímicas. Depois eles receberam o Prêmio Nobel por seu estudo, que marca o começo de toda a Biologia Molecular. Beadle e Tatum fizeram seu trabalho no fungo haplóide *Neurospora*, que já vimos em capítulos anteriores. Seu plano era investigar o controle genético da química celular. No que se tornou o enfoque genético padrão direto, eles primeiro irradiaram as células de *Neurospora* para produzir mutações e, então, testaram culturas crescidas dos ascósporos quanto a fenótipos mutantes de interesse. Eles descobriram numerosos mutantes que tinham nutrição defectiva. Especificamente, eles eram mutantes auxotróficos, do tipo descrito para bactérias no Capítulo 5. Tais mutantes requerem o suprimento de um nutriente que um fungo tipo selvagem é capaz de produzir por si mesmo, sugerindo que o mutante era defectivo para alguma etapa sintética normal. Beadle e Tatum deduziram que cada mutação que gerava a necessidade de um nutriente era herdada como uma mutação em um só gene, porque cada uma dava uma proporção de 1:1 quando cruzada com um tipo selvagem. Representemos a mutação auxotrófica por *aux*,

$$\begin{array}{c} + \times aux \\ \downarrow \\ \text{prole} \quad \frac{1}{2} + \\ \quad \quad \frac{1}{2} aux \end{array}$$

A Figura 6.11 mostra o procedimento experimental usado por Beadle e Tatum. Sua análise enfocou um grupo particular de linhagens mutantes que precisam do aminoácido arginina para crescer (os mutantes eram auxotróficos para arginina). Eles descobriram que essas mutações auxotróficas estavam mapeadas em três loci diferentes em três cromossomos separados. Vamos chamar os genes nos três loci de *arg-1*, *arg-2* e *arg-3*. Uma importante conquista foi sua descoberta de que os auxotróficos para cada um dos três loci diferiam em sua resposta aos compostos estruturalmente correlatos ornitina e citrulina (Figura 6.12). Os mutantes *arg-1* cresciam quando suplementados por ornitina, citrulina ou arginina. Os mutantes *arg-2* cresciam quando recebiam arginina ou citrulina, mas não ornitina. Os mutantes *arg-3* só cresciam quando era dada arginina. Esses resultados estão resumidos no Quadro 6.1.

As enzimas celulares já eram conhecidas interconvertendo tais compostos correlatos. Com base nas propriedades dos mutantes *arg*, Beadle e Tatum *et al.* propuseram uma via bioquímica para tais conversões em *Neurospora*:



Essa via explica perfeitamente as três classes de mutantes mostradas no Quadro 6.1. Nesse modelo, os mutantes *arg-1* têm uma falta da enzima X, e, assim, são incapazes de converter o precursor em ornitina como primeira etapa na produção de arginina. Entretanto, eles têm enzimas normais Y e Z, e, assim, os mutantes *arg-1* são capazes de produzir arginina se

Descobridor dos erros hereditários do metabolismo

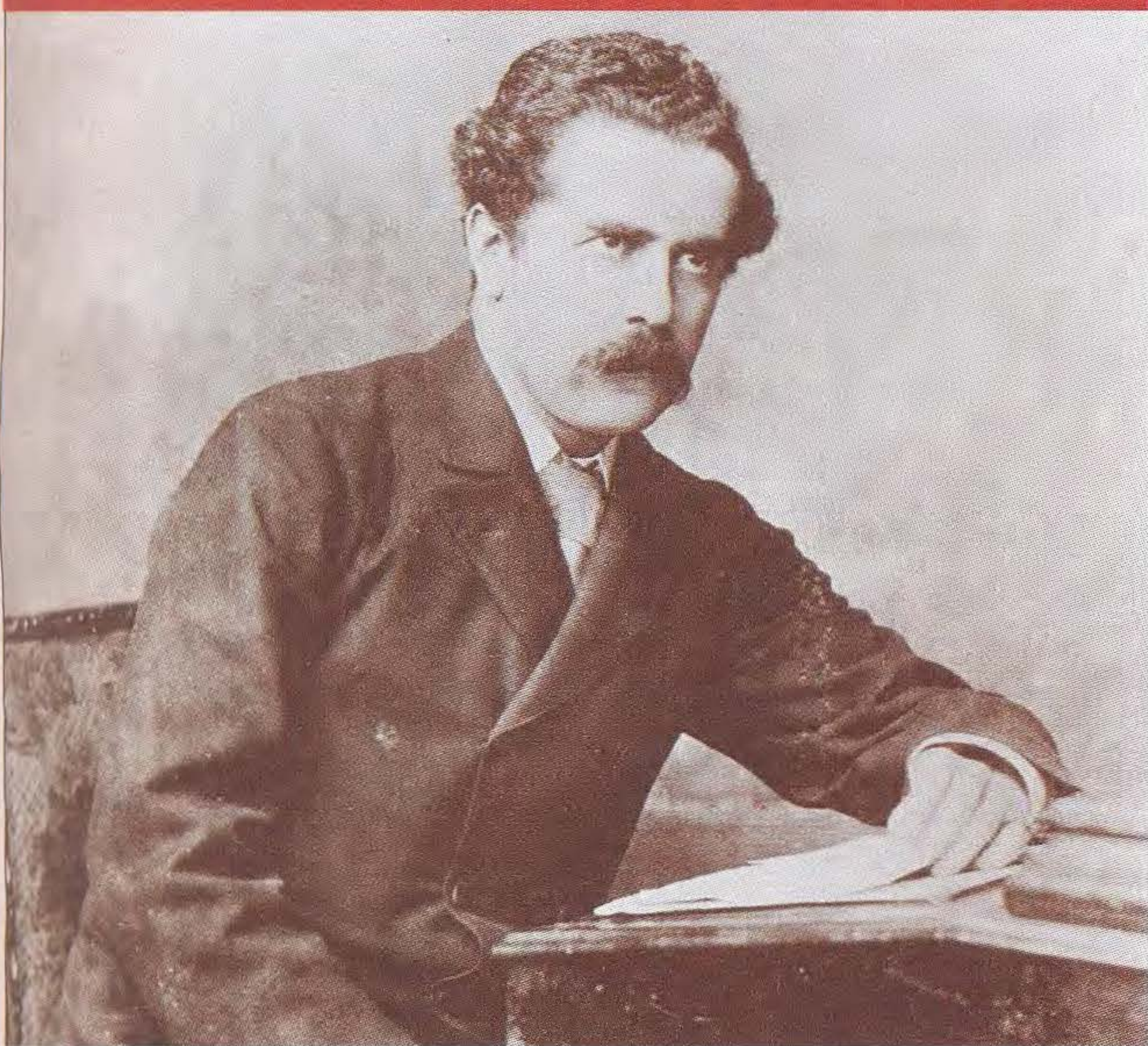


FIG. 6.10 O médico britânico Archibald Garrod (1857–1936).

Isolamento de uma mutação auxotrófica para arginina

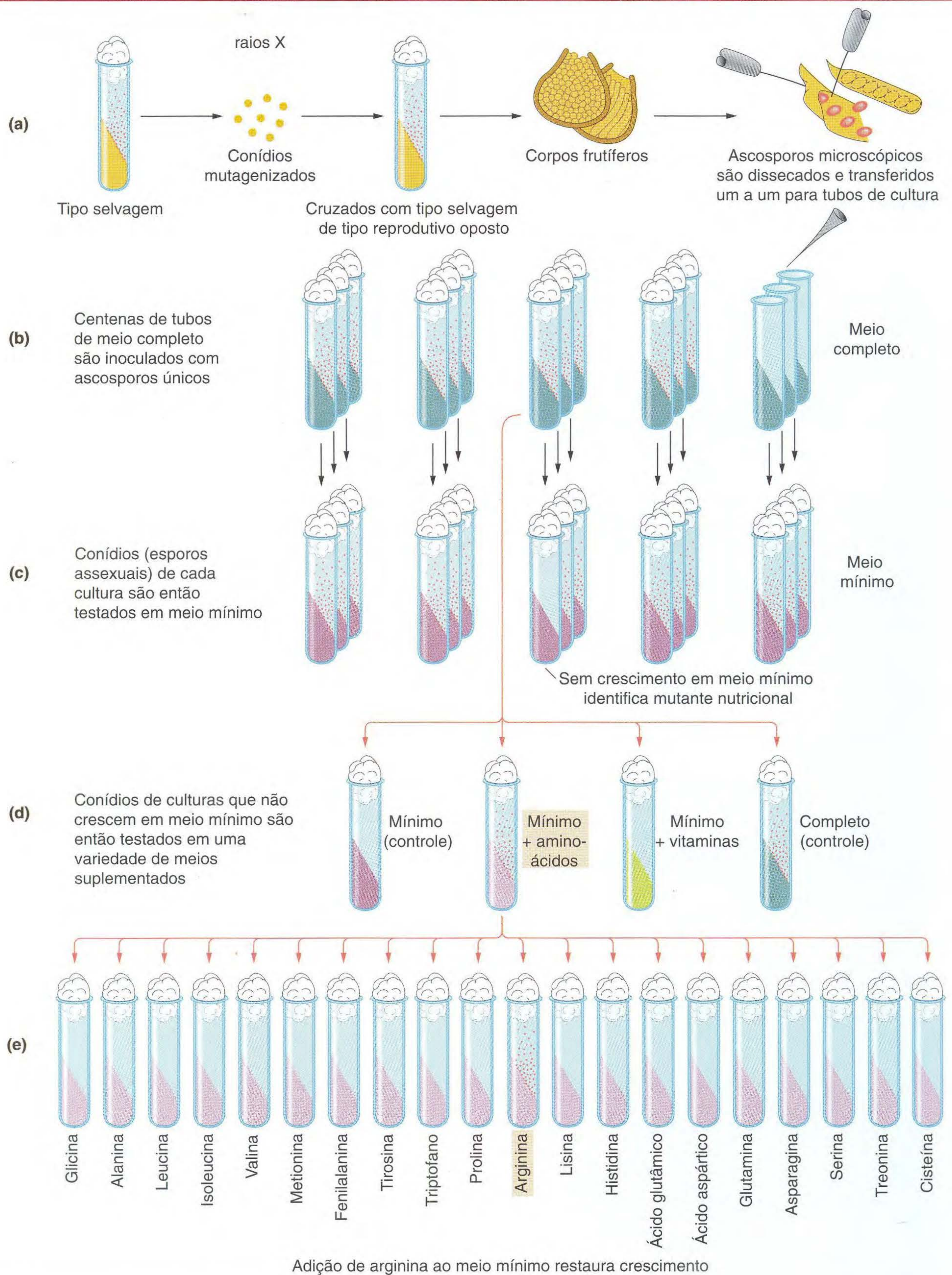


FIG. 6.11 Enfoque experimental usado por Beadle e Tatum para gerar grandes números de mutantes em *Neurospora*. Essa análise particular mostra o isolamento de um mutante *arg⁻*. [De P.J. Russell, *Genetics*, 2nd ed. Scott, Foresman.]

Arginina e seus análogos

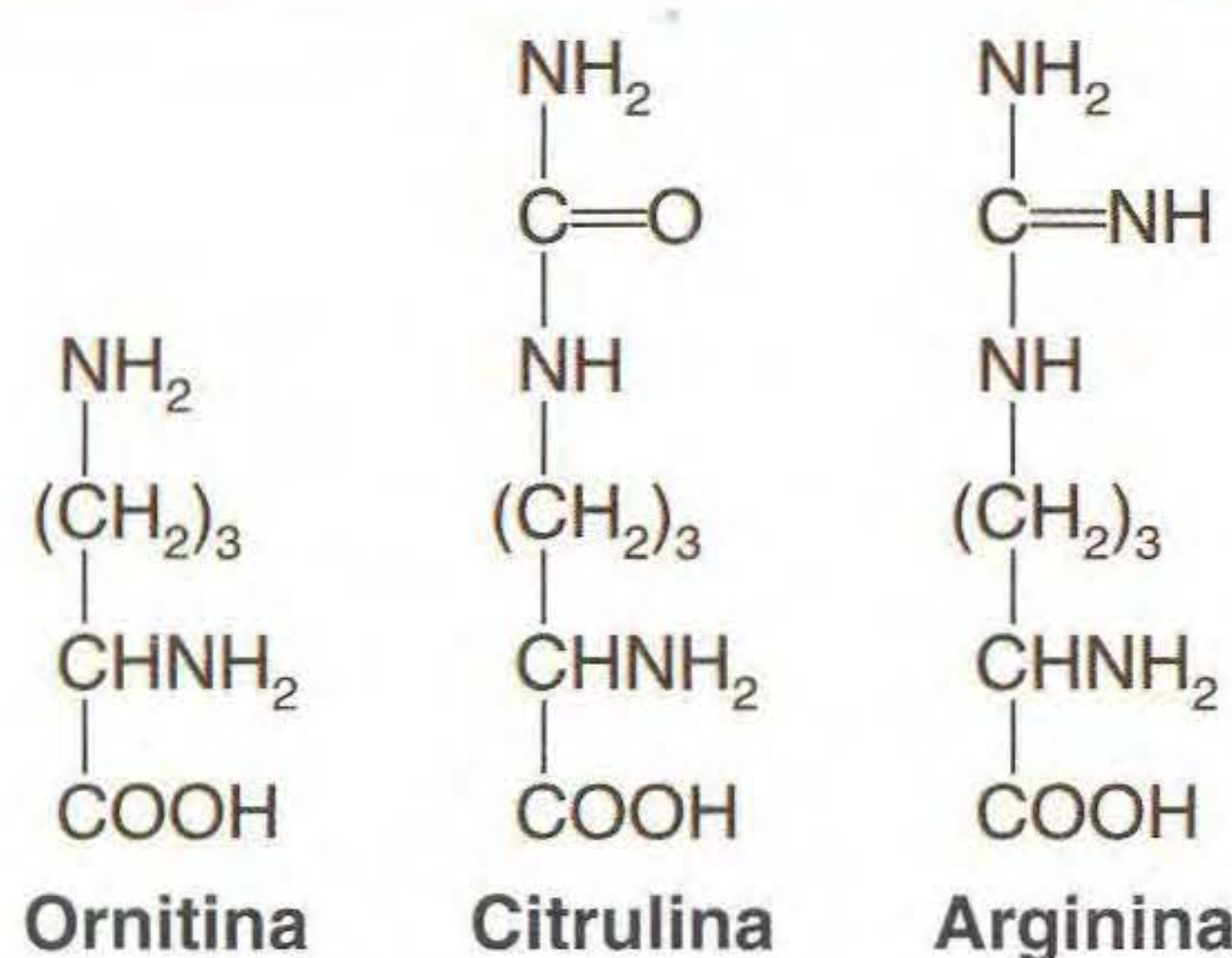
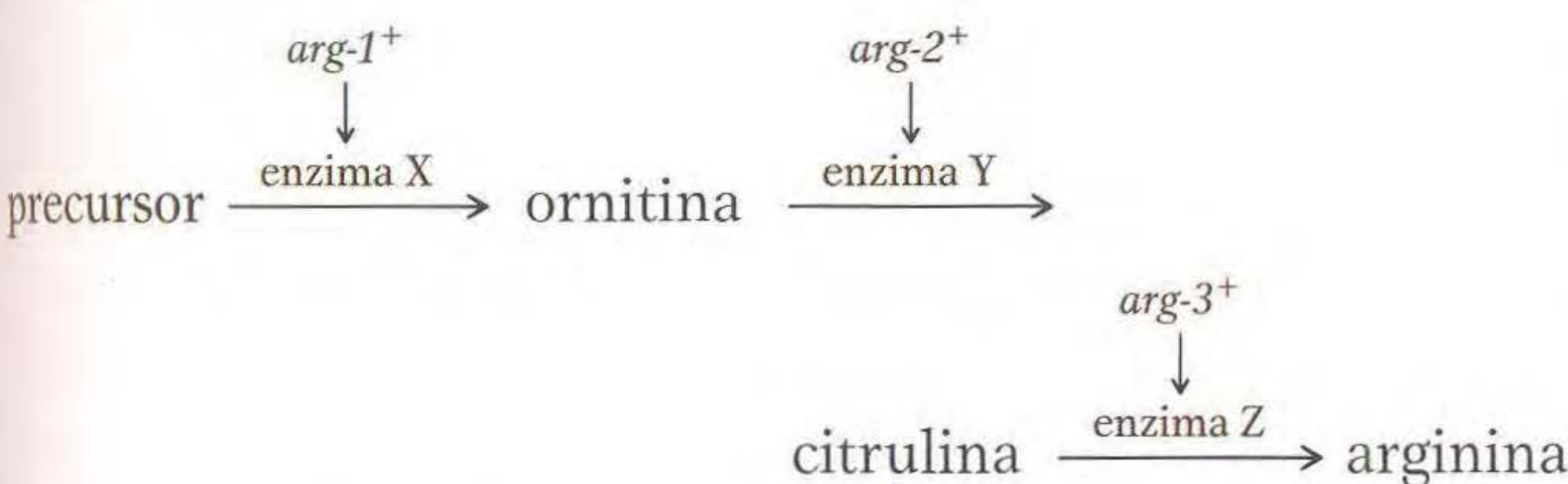


FIG. 6.12 As estruturas químicas da arginina e os compostos estruturalmente correlatos citrulina e ornitina.

receberem ornitina ou citrulina. Similarmente, os mutantes *arg-2* não têm a enzima Y, e os mutantes *arg-3* não têm a enzima Z. Assim, uma mutação em determinado gene supostamente interfere na produção de uma única enzima. A falta da enzima cria um bloqueio em alguma via biossintética. O bloqueio pode ser evitado dando às células qualquer composto que normalmente vem após o bloqueio na via.

Agora podemos diagramar um modelo bioquímico mais completo:



Esse modelo brilhante, que foi inicialmente conhecido como *hipótese um gene – uma enzima*, foi a fonte da primeira compreensão das funções dos genes: os genes de algum modo são responsáveis pela função das enzimas, e cada gene aparentemente controla uma enzima específica em uma série de etapas interconectadas em uma via bioquímica. Outros pesquisadores obtiveram resultados similares para outras vias biossintéticas, e a hipótese logo teve aceitação geral. Todas as proteínas, sejam enzimas ou não, também são codificadas por genes, e, assim, a frase foi redefinida para tornar-se **hipótese um gene – um polipeptídeo**.* (Lembre que um polipeptídeo é o tipo mais simples de proteína, uma única cadeia de aminoácidos.) Logo ficou claro que um gene codifica a *estrutura física* de uma proteína, que, por sua vez, determina sua função. A hipótese de Beadle e Tatum tornou-se um dos grandes conceitos unificantes em biologia, pois forneceu uma ponte que liga as duas importantes áreas de pesquisa da genética e bioquímica.

(Entre parênteses, devemos acrescentar que, embora a grande maioria dos genes codifiquem proteínas, alguns codificam RNA que têm funções especiais. Todos os genes são transcritos para fazer RNA. Os genes codificantes de proteínas são transcritos em RNA mensageiro (mRNA), que então

*N.R.T.: Atualmente modificada para um gene (ou cístron) um ou mais tipos de polipeptídeos.

Quadro 6.1 Crescimento de Mutantes *arg* em Resposta a Suplementos

Mutante	Suplemento		
	Ornitina	Citrulina	Arginina
<i>arg-1</i>	+	+	+
<i>arg-2</i>	-	+	+
<i>arg-3</i>	-	-	+

Nota: Um sinal + significa crescimento; um sinal - significa sem crescimento.

é traduzido em proteína. Entretanto, o RNA codificado por uma minoria de genes nunca é traduzido em proteína, porque o próprio RNA tem uma função única. Devemos chamá-los de **RNA funcionais**. Alguns exemplos são os RNA transportadores, RNA ribossômicos e pequenos RNA citoplasmáticos. Veremos mais sobre eles em capítulos posteriores.)

A síntese química nas células é por vias de etapas sequenciais catalisadas por enzimas. Os genes que codificam as enzimas de uma via específica constituem um subgrupo funcionalmente interagente do genoma.

Interação gênica em outros tipos de vias

A noção de que os genes interagem por meio de vias é poderosa e tem aplicação em todos os organismos. A via de arginina de *Neurospora* é um exemplo de uma via sintética, uma cadeia de conversões enzimáticas que produz nutrientes essenciais. Podemos ampliar a idéia novamente para outro caso já introduzido, a doença fenilcetonúria (PKU), que é causada por um alelo autossômico recessivo. A PKU resulta da incapacidade de converter fenilalanina em tirosina. Conseqüentemente, a fenilalanina acumula-se e é espontaneamente convertida em um composto tóxico, o ácido fenilpirúvico. O gene de PKU é parte de uma via sintética, como a via de arginina de *Neurospora*, e parte dela é mostrada na Figura 6.13. A ilustração inclui várias outras doenças causadas por bloqueios em etapas nessa via (incluindo alcaptonúria, a doença investigada por Garrod).

Além de uma via puramente intracelular de síntese, uma ampliação da via bioquímica no sistema de expressão de PKU envolve o movimento de produtos químicos pelo corpo. A Figura 6.14 mostra que existem muitas etapas na via levando desde a ingestão da fenilalanina dietética até o fenótipo de PKU de prejuízo do desenvolvimento cognitivo. Qualquer uma delas pode ser controlada por genes. Primeiro, a quantidade de fenilalanina na dieta é de suma importância, pois, como descrito no Capítulo 2, a restrição de fenilalanina pode aliviar os sintomas das pessoas homozigotas para o mutante. Em seguida, a fenilalanina deve ser transportada para locais apropriados no fígado, a “fábrica química” do corpo. No fígado, PAH deve agir em conjunto com seu co-fator, a tetraidrobiopterina. Podemos deduzir que as mutações nos genes que não PAH podem causar níveis elevados de fenilalanina. Um gene necessário para a síntese de tetraidrobiopterina é um

Uma via sintética e doenças associadas

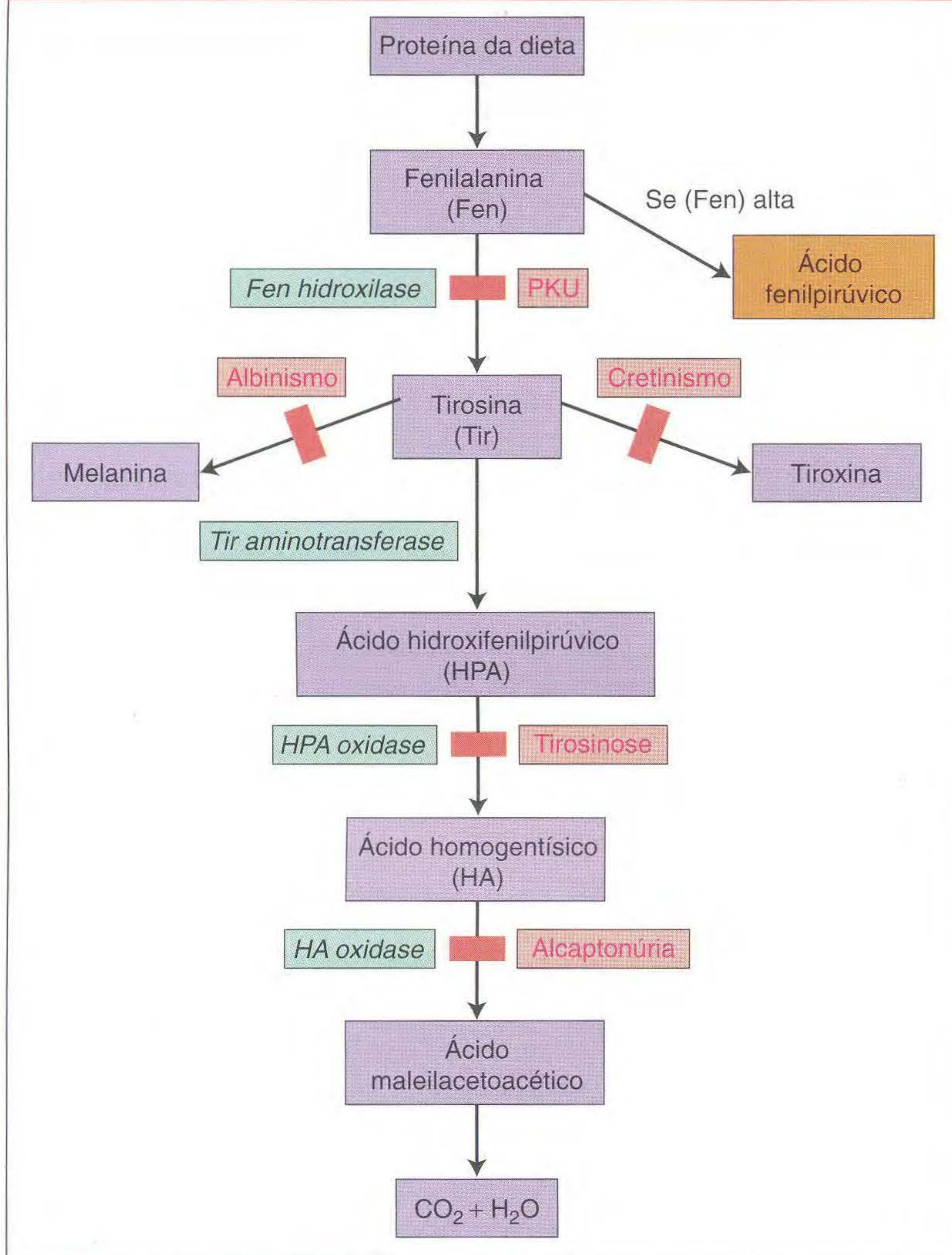


FIG. 6.13 Um trecho da via metabólica de fenilalanina em humanos, incluindo doenças associadas aos bloqueios enzimáticos. A doença PKU é produzida quando a enzima fenilalanina hidroxilase tem uma disfunção. O acúmulo de fenilalanina resulta em um aumento de ácido fenilpirúvico, que interfere no desenvolvimento do sistema nervoso. [De I.M. Lerner and W.J. Libby, *Heredity, Evolution, and Society*, 2nd ed. Copyright 1976 by W.H. Freeman and Company.]

desses genes, e um gene necessário para expressar o transportador que leva fenilalanina para o fígado é outro. Podemos também ver como as mutações em outros locais podem evitar que pessoas com PKU tenham desenvolvimento cognitivo anormal. Se um nível anormalmente alto de ácido fenilpirúvico for produzido, para afetar o desenvolvimento cognitivo ele deve ser transportado para o cérebro na corrente sanguínea e, então, passar pela barreira hematoencefálica. Dentro do cérebro, os processos desenvolvimentais devem ser suscetíveis à ação detrimental do ácido fenilpirúvico. Cada uma das etapas na via corpórea é um sítio possível no qual os alelos mutantes (ou variação ambiental) podem ser expressos. A interação de genes na cor dos cabelos humanos, já mencionada, mostra uma complexidade similar de interação gênica em vias dentro e dentre células do corpo.

As vias de transdução de sinal foram resumidamente descritas no começo deste capítulo. Uma das vias de transdução de sinal mais bem compreendida foi desenvolvida de uma análise genética da resposta reprodutiva de levedura. Lembre que dois tipos reprodutivos, determinados pelos alelos MATa e MAT α , são necessários para que ocorra reprodução em levedura. Quando uma célula está na presença de outra célula do tipo reprodutivo oposto, ela sofre uma série de mudanças de forma e comportamento para se reproduzir. A resposta reprodutiva é disparada por uma via de transdução de sinal que requer a ação sequencial de um conjunto de genes. Esse conjunto de genes foi descoberto por uma análise de interação padrão de mutantes com resposta reprodutiva aberrante (a maioria era estéril). As etapas foram juntadas usando os enfoques da seção seguinte.

O sinal que move as coisas é um ferormônio (hormônio) reprodutivo liberado pelo tipo reprodutivo oposto; o ferormônio liga-se a um receptor de membrana, que é acoplado a uma proteína G dentro da membrana e ativa a proteína. A proteína G, por sua vez, move uma série de fosforilações sequenciais de proteína chamadas de cascata de cinases. Finalmente, a cascata ativa a transcrição de um conjunto de genes específicos de reprodução que permitem que a célula se reproduza. Uma mutação em qualquer uma dessas etapas pode perturbar o processo reprodutivo.

As vias desenvolvimentais compreendem as etapas pelas quais um zigoto torna-se um organismo adulto. Essas etapas podem incluir a regulação gênica, a transdução de sinal, a diferenciação dos tecidos e o movimento. As vias desenvolvimentais serão vistas em detalhe no Capítulo 12, mas a interação de genes nessas vias é analisada do mesmo modo, como veremos em seguida.

► 6.3 Interações Gênicas Interferentes

Como efetuar uma análise genética que revela os genes que interagem para contribuir para uma propriedade biológica determinada já foi descrito no começo deste capítulo. Resumidamente, o enfoque é o seguinte:

Etapa 1. Tratar as células com agentes causadores de mutação (mutágenos), tais como radiação ultravioleta. Esse tratamento produz um grande conjunto de mutantes com uma expressão anormal da propriedade em estudo. Cada mutante é verificado para confirmar a herança monogênica (proporções mendelianas monoíbridas). É avaliado o efeito fenotípico de cada mutante no desenvolvimento da propriedade.

Algumas vias na PKU

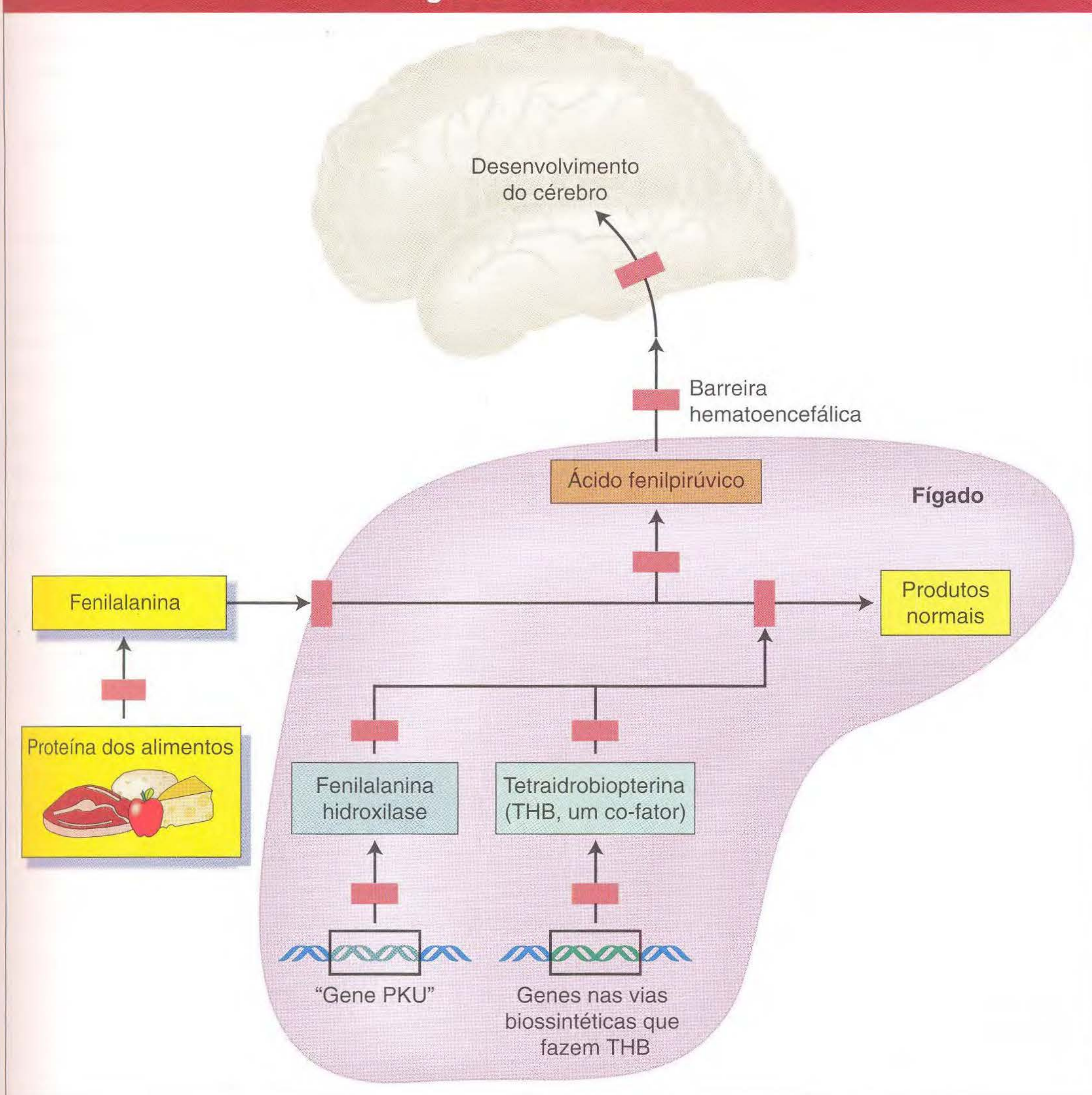


FIG. 6.14 A expressão da doença PKU envolve uma série complexa de etapas. Os retângulos vermelhos indicam etapas onde a variação ou bloqueio é possível.

Etapa 2. Testar esses mutantes para determinar o número total de loci gênicos que tomam parte e quais mutações são alelos do mesmo gene.

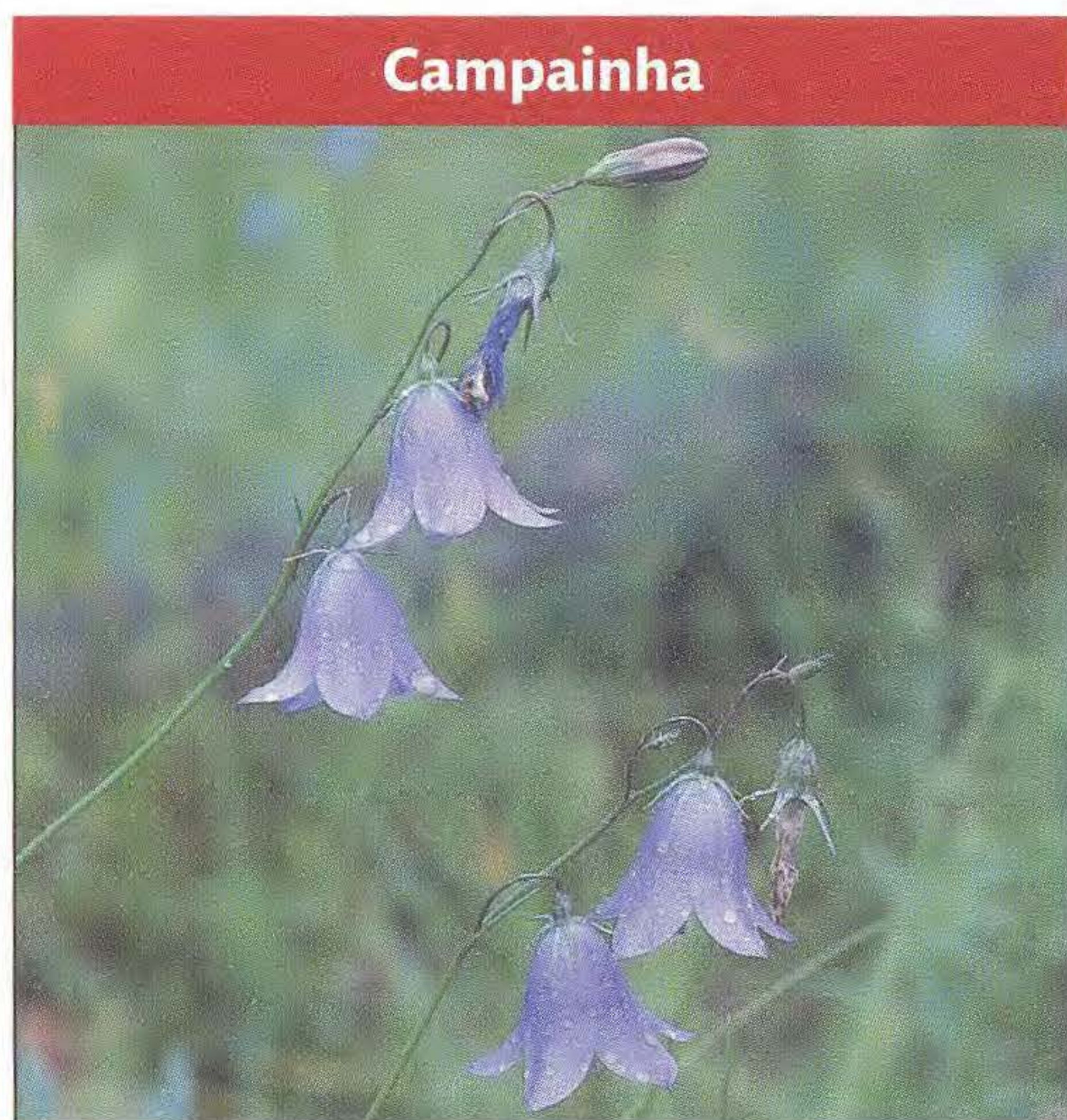
Etapa 3. Combine as mutações par a par fazendo cruzamentos para formar **mutantes duplos** para ver se eles interagem. A interação gênica é deduzida do fenótipo do mutante duplo, essencialmente qualquer desvio da combinação simples de fenótipos monogênicos mutantes. Se os alelos mutantes interagem, então podemos deduzir que os genes tipo selvagem também interagem normalmente.

Um procedimento que deve ser feito antes de testar as interações é determinar se cada mutação é um locus diferente. A triagem de mutantes pode não ter favorecido intencionalmente alguns genes. Nesse caso, em uma coleção de, digamos, 60 mutações independentes, apenas 20 loci podem estar representados, cada um com cerca de três mutações. Assim, o conjunto de loci gênicos precisa ser definido, como mostrado na seção seguinte.

Definição de conjuntos de genes usando o teste de complementação

Como é possível decidir se duas mutações pertencem ao mesmo gene? Existem vários modos. Primeiro, cada alelo mutante pode ser mapeado. Então, se duas mutações estão mapeadas em dois loci cromossômicos diferentes, provavelmente eles são de genes diferentes. Entretanto, esse enfoque consome tempo em um grande grupo de mutações. O enfoque geralmente usado é o **teste de complementação**. Embora o teste de complementação possa ser aplicado apenas a mutações recessivas, ele é relativamente simples e rápido de usar, e essas características respondem por sua popularidade.

Em um organismo diplóide, a essência do teste de complementação é cruzar duas linhagens mutantes homozigotas para produzir um heterozigoto. Se o heterozigoto é tipo selvagem, então as duas mutações devem ter estado em genes diferentes. Nesse caso, as duas mutações são ditas tendo se *complementado*. Ilustremos o teste de complementação com um exemplo das plantas campainha (gênero *Campanula*). A



Flores da planta campainha (espécie *Campanula*). [Gregory G. Dimijian/Photo Researchers.]

cor da flor do tipo selvagem dessa planta é azul. Suponhamos que, em uma caça ao mutante, obtivemos três mutantes de pétalas brancas e eles estão disponíveis como linhagens puras homozigotas. Todas parecem iguais, e, assim, não sabemos *a priori* se são geneticamente idênticas. Chamaremos as linhagens mutantes de \$, £ e ¥, para evitar qualquer símbolo usando letras que possam significar dominância. Quando os mutantes são cruzados com tipo selvagem, cada mutante dá os mesmos resultados na F_1 e F_2 , como o seguinte:

branca \$ \times azul $\rightarrow F_1$, toda azul $\rightarrow F_2$, $\frac{3}{4}$ azul, $\frac{1}{4}$ branca

branca £ \times azul $\rightarrow F_1$, toda azul $\rightarrow F_2$, $\frac{3}{4}$ azul, $\frac{1}{4}$ branca

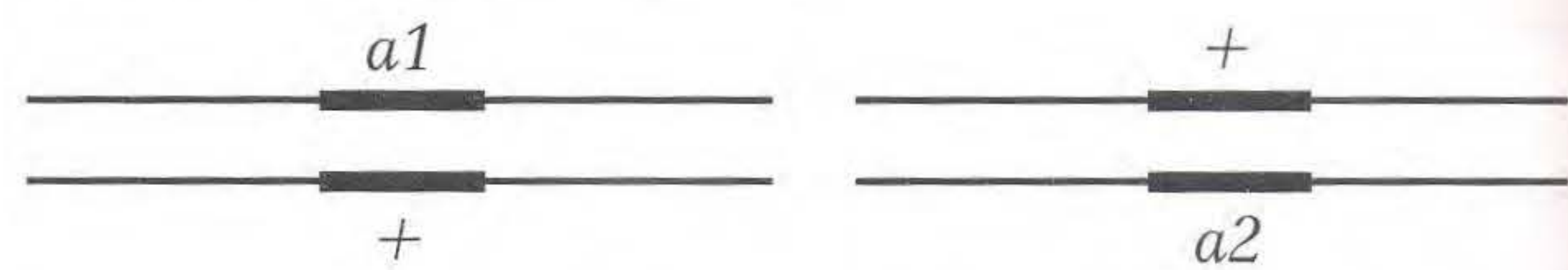
branca ¥ \times azul $\rightarrow F_1$, toda azul $\rightarrow F_2$, $\frac{3}{4}$ azul, $\frac{1}{4}$ branca

Em cada caso, os resultados mostram que a condição mutante é determinada pelo alelo recessivo de um único gene. Entretanto, existem três alelos de *um* gene, de dois genes ou de três genes? Como os mutantes são recessivos, a pergunta pode ser respondida com o teste de complementação, que indaga se os mutantes se *complementam* uns aos outros. A nível operacional, a complementação é definida do seguinte modo.

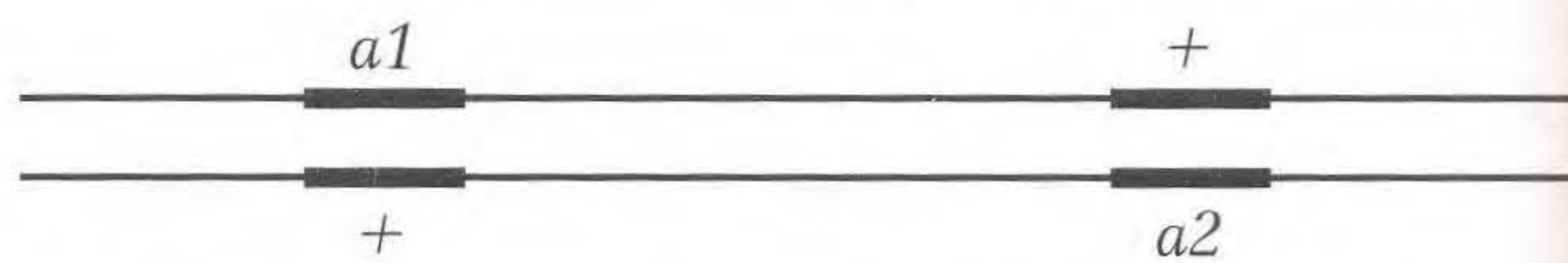
Complementação é a produção de um fenótipo tipo selvagem quando dois genomas haplóides, tendo mutações recessivas diferentes, são unidos na mesma célula.

Em uma diplóide, o teste de complementação é feito entrecruzando dois indivíduos que são homozigotos para mutações recessivas diferentes. A etapa seguinte é observar se a prole tem o fenótipo tipo selvagem. Se a prole é toda tipo selvagem, as duas mutações recessivas devem estar em genes *diferentes*, pois os respectivos alelos tipo selvagem dão a função tipo selvagem. Aqui, chamaremos os genes de $a1$ e $a2$, para seus alelos mutantes. Podemos representar os heterozigotos como segue, dependendo de se os genes estão no mesmo cromossomo ou em cromossomos diferentes:

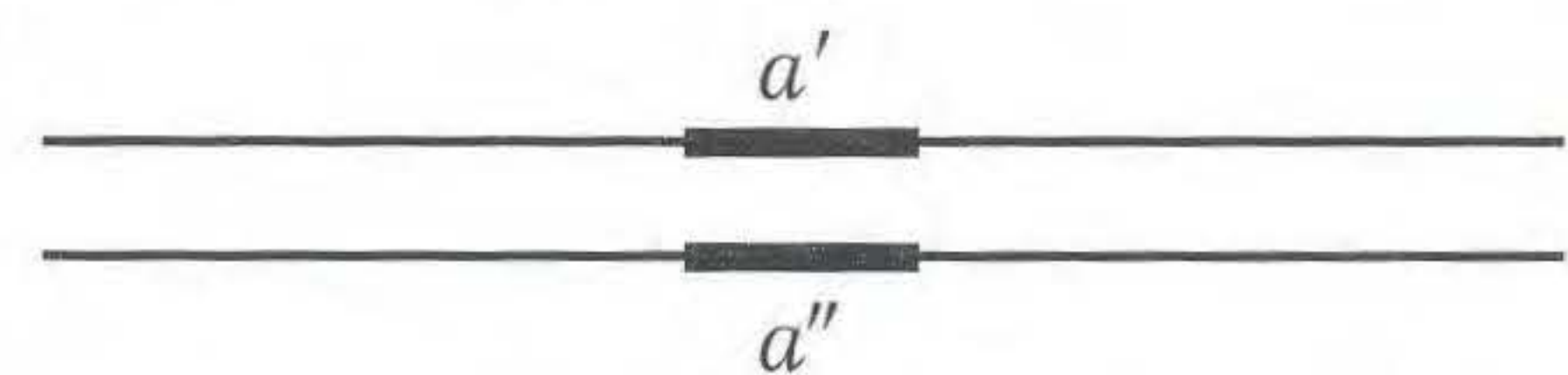
Cromossomos diferentes



Mesmo cromossomo (mostrado em configuração trans)



Entretanto, se a prole *não* é tipo selvagem, então as mutações recessivas devem ser alelos do mesmo gene. Como ambos os alelos do gene são mutantes, não há alelo tipo selvagem para dar a função tipo selvagem. Tais alelos podem ser tidos, em geral, como a' e a'' , usando primos para distingui-los entre dois alelos mutantes diferentes de um gene cujo alelo tipo selvagem é a^+ . Esses alelos podem ter sítios mutantes diferentes dentro do mesmo gene, mas ambos serão não-funcionais. O heterozigoto a'/a'' seria



Voltemos ao exemplo de campainha e ao entrecruzamento dos mutantes para testar a complementação. Suponha que os resultados dos entrecruzamentos dos mutantes \$, £ e ¥ sejam

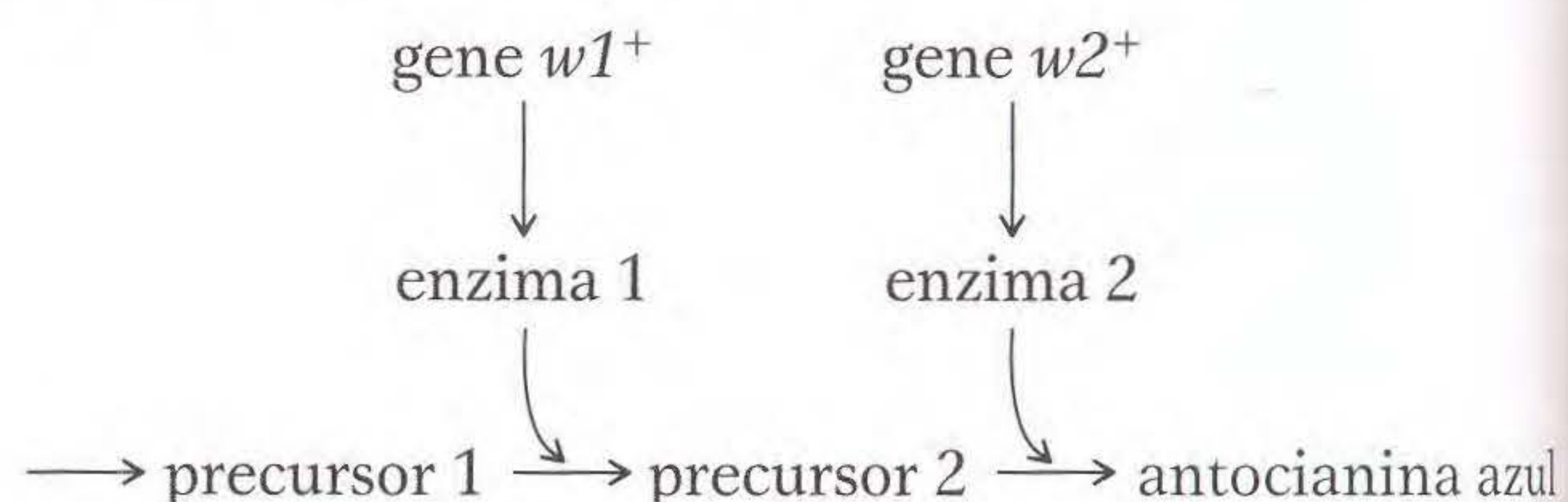
branca \$ \times branca £ $\rightarrow F_1$, toda branca

branca \$ \times branca ¥ $\rightarrow F_1$, toda azul

branca £ \times branca ¥ $\rightarrow F_1$, toda azul

Por esse conjunto de resultados, podemos concluir que os mutantes \$ e £ devem ser causados por alelos de um gene (digamos, $w1$) porque eles não se complementam; mas ¥ deve ser causado por um alelo mutante de outro gene ($w2$), pois é vista a complementação.

Como a complementação funciona a nível molecular? A cor azul normal da flor de campainha é causada por um pigmento azul chamado *antocianina*. Os pigmentos são substâncias que absorvem algumas cores da luz; com relação à campainha, a antocianina absorve todos os comprimentos de onda exceto azul, que é refletido no olho do observador. Entretanto, essa antocianina é feita de precursores químicos que não são pigmentos; isto é, eles não absorvem luz de nenhum comprimento de onda específico e, simplesmente, refletem de volta a luz branca do sol para o observador, dando um aspecto branco. O pigmento azul é o produto final de uma série de conversões bioquímicas de não-pigmentos. Cada etapa é catalisada por uma enzima específica codificada por um gene específico. Podemos explicar os resultados com uma via conforme o seguinte:



Uma mutação homozigota em um dos genes levará a um acúmulo de um precursor que simplesmente faz a planta branca. Agora as designações mutantes podem ser escritas do seguinte modo:

$$\begin{array}{ll} \$ & w1_{\$}/w1_{\$} \cdot w2^{+}/w2^{+} \\ \pounds & w1_{\pounds}/w1_{\pounds} \cdot w2^{+}/w2^{+} \\ \text{¥} & w1^{+}/w1^{+} \cdot w2_{\text{¥}}/w2_{\text{¥}} \end{array}$$

Entretanto, na prática, os símbolos de subscritos seriam abandonados e os genótipos seriam escritos do seguinte modo:

$$\begin{array}{ll} \$ & w1/w1 \cdot w2^{+}/w2^{+} \\ \pounds & w1/w1 \cdot w2^{+}/w2^{+} \\ \text{¥} & w1^{+}/w1^{+} \cdot w2/w2 \end{array}$$

Assim, uma F_1 de $\$ \times \pounds$ será

$$w1/w1 \cdot w2^{+}/w2^{+}$$

Essas plantas de F_1 terão dois alelos defectivos para $w1$ e serão, portanto, bloqueadas na etapa 1. Muito embora a enzima 2 seja totalmente funcional, ela não tem substrato no qual agir; assim, nenhum pigmento azul será produzido, e o fenótipo será branco.

As plantas de F_1 de outros cruzamentos, entretanto, terão os alelos tipo selvagem para ambas as enzimas necessárias para levar os intermediários até o produto final azul. Seus genótipos serão

$$w1^{+}/w1 \cdot w2^{+}/w2$$

Assim, vemos que a complementação é de fato um resultado da interação cooperativa dos alelos *tipo selvagem* dos dois genes. A Figura 6.15 resume a interação da complementação e não-complementação de mutantes brancos nos níveis genético e celular.

Em um organismo haplóide, o teste de complementação não pode ser feito por entrecruzamento. Em fungos, um modo alternativo junta alelos mutantes para testar a complementação: a fusão resultando em um **heterocácion** (Figura 6.16). As células dos fungos fundem-se prontamente. Quando duas linhagens diferentes se fundem, os núcleos haplóides de linhagens diferentes ocupam uma célula, que é o *heterocácion* (do grego, núcleos diferentes). Os núcleos em um heterocácion geralmente não se fundem. De certo modo, essa condição é uma “imitação” diplóide.

Suponha que, em linhagens diferentes, existam mutações em dois genes diferentes conferindo o mesmo fenótipo mutante — por exemplo, necessidade de arginina. Chamaremos esses genes de *arg-1* e *arg-2*. Esses genótipos das duas linhagens podem ser representados como *arg-1* · *arg-2*⁺ e *arg-1*⁺ · *arg-2*. Essas duas linhagens podem ser fundidas para formar um heterocácion com os dois núcleos em um citoplasma compartilhado:

$$\begin{array}{l} \text{Núcleo 1 é } arg-1 \cdot arg-2^{+} \\ \text{Núcleo 2 é } arg-1^{+} \cdot arg-2 \end{array}$$

Como os produtos gênicos são feitos em um citoplasma comum, os dois alelos tipo selvagem podem exercer seu efei-

to dominante e cooperar para produzir um heterocácion de fenótipo tipo selvagem. Em outras palavras, as duas mutações se complementam, como fariam em um diplóide. Se as mutações fossem alelos do mesmo gene, não haveria complementação.

Quando dois alelos mutantes recessivos derivados independentemente produzindo fenótipos recessivos similares falham em se complementar, eles devem ser alelos do mesmo gene.

Análise de mutantes duplos por mutações aleatórias

Lembre que, para saber se dois genes interagem, precisamos avaliar o fenótipo do mutante duplo para ver se é outro que não a combinação de mutações únicas. O mutante duplo é obtido por entrecruzamento. A F_1 é obtida como parte do teste de complementação. Assim, com a suposição de que foi observada complementação, sugerindo genes diferentes, a F_1 é autofecundada para se obter uma F_2 . A F_2 deve conter o mutante duplo. O mutante duplo pode então ser identificado observando-se as proporções mendelianas. Por exemplo, se é obtida uma proporção mendeliana 9:3:3:1, o fenótipo presente apenas em 1/16 da prole representa o mutante duplo (o “1” na 9:3:3:1). Nos casos de interação gênica, entretanto, o fenótipo do mutante duplo pode não ser distinto, mas corresponderá a um dos mutantes únicos. Nesse caso, resultará uma proporção mendeliana modificada, tal como 9:3:4 ou 9:7.

A proporção mendeliana padrão 9:3:3:1 é o caso mais simples, esperada se não houver interação gênica e se duas mutações testadas são em cromossomos diferentes. Essa proporção 9:3:3:1 é a hipótese nula: qualquer proporção mendeliana modificada que representa um desvio dessa hipótese nula seria informativa, como o exemplo seguinte mostrará.

A proporção 9:3:3:1: sem interação gênica Como linha básica, comecemos com o caso no qual dois genes mutantes não interagem, uma situação em que esperamos a proporção 9:3:3:1. Vejamos a herança da cor da pele nas cobras de milharal. A cor natural das cobras é um padrão de camuflagem preto e laranja, como mostrado na Figura 6.17a. O fenótipo é produzido por dois pigmentos separados, ambos sob controle genético. Um gene determina o pigmento laranja, e os alelos que vamos considerar são o^{+} (presença de pigmento laranja) e o (ausência de pigmento laranja). Um outro gene determina o pigmento preto, e seus alelos são b^{+} (presença do pigmento preto) e b (ausência do pigmento preto). Esses dois genes não estão ligados. O padrão natural é produzido pelo genótipo $o^{+}/- ; b^{+}/-$. Uma cobra que é $o/o ; b^{+}/-$ é preta porque não tem o pigmento laranja (Figura 6.17b), e uma cobra que é $o^{+}/- ; b/b$ é laranja porque não

A base molecular da complementação genética

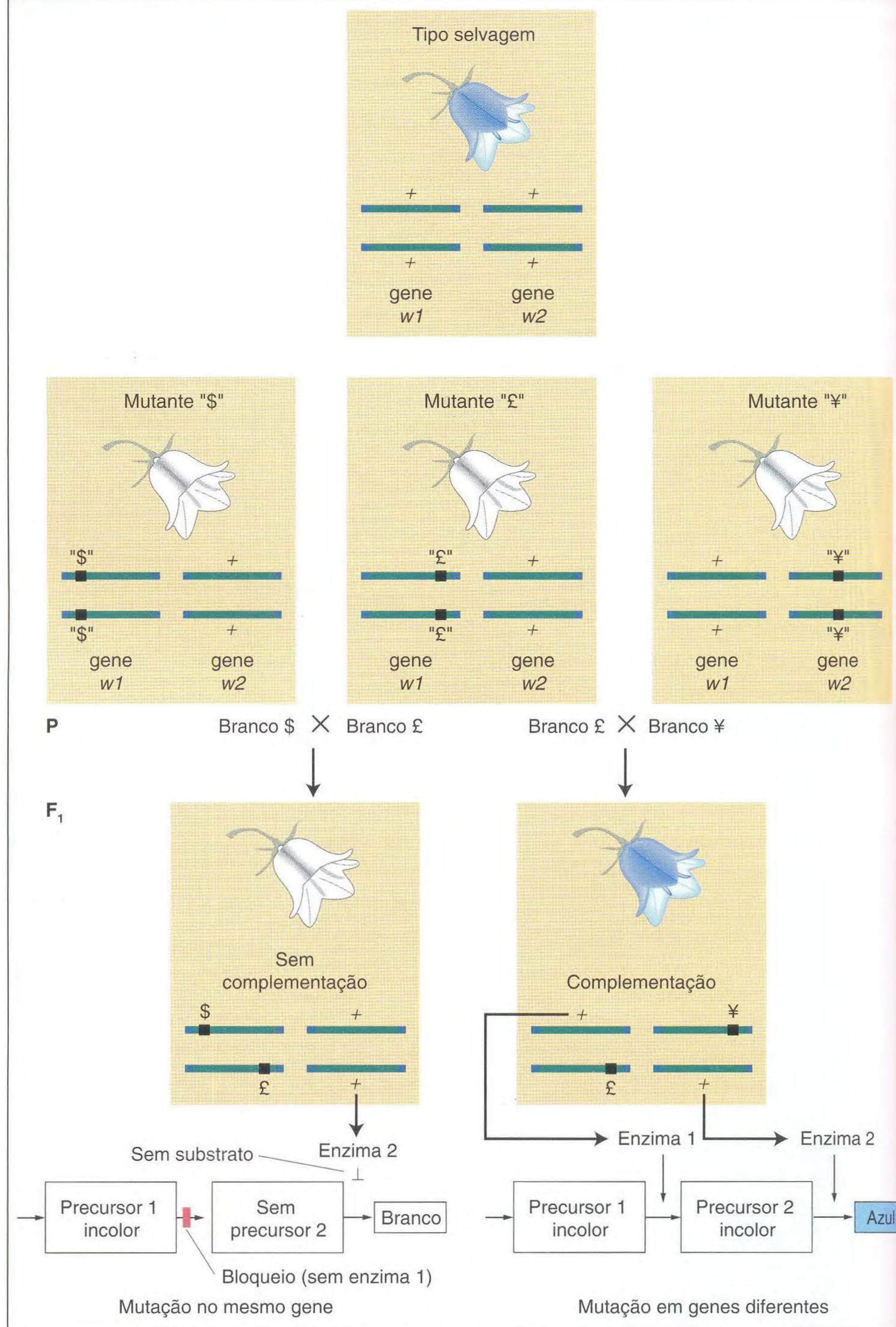


FIG. 6.15 Esses três mutantes brancos fenotipicamente idênticos de campainha, £, ¥ e \$, são entrecruzados. As mutações no mesmo gene (tal como £ e \$) não se complementam, pois a F_1 tem um gene com dois alelos mutantes. A via é bloqueada e as flores são brancas. Quando as mutações são em genes diferentes (tais como £ e ¥), há complementação dos alelos tipo selvagem de cada gene no heterozigoto de F_1 . O pigmento é sintetizado e as flores são azuis. (Que resultado você poderia prever para o cruzamento de \$ e ¥?)

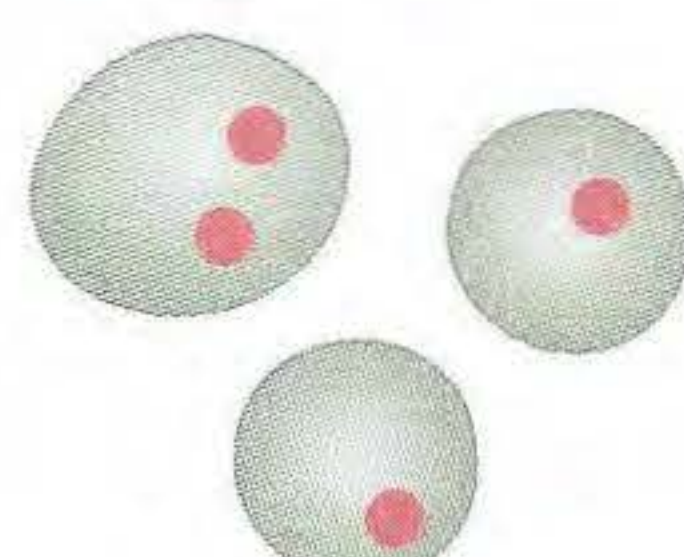
tem o pigmento preto (Figura 6.17c). O duplo homozigoto recessivo o/o ; b/b é albino (Figura 6.17d). Note, entretanto, que a cor rosa-claro do albino é por um outro pigmento: a hemoglobina do sangue, que é visível através da pele dessa cobra quando os outros pigmentos estão ausentes. A cobra albina mostra claramente que há um outro elemento para o padrão de pigmentação da pele em adição ao pigmento: o motivo repetido no e ao redor do qual o pigmento é depositado.

Se uma cobra homozigota laranja e uma homozigota preta são cruzadas, a F_1 é tipo selvagem (camuflada), demonstrando complementação:

$$\begin{array}{ccc}
 \text{♀ } o^+/o^+ ; b/b & \times & \text{♂ } o/o ; b^+/b^+ \\
 \text{(laranja)} & & \text{(preto)} \\
 & \downarrow & \\
 F_1 & o^+/o ; b^+/b & \\
 & \text{(camuflagem)} &
 \end{array}$$

Teste de complementação usando um heterocácion

células *arg-1*, defectivas para uma enzima específica na via de síntese de arginina



células *arg-2*, defectivas para uma enzima diferente na via de síntese de arginina

Fusão



Heterocácion cresce sem arginina

FIG. 6.16 Um heterocácion de *Neurospora* e fungos similares mimetizam um estado diplóide. Quando as células vegetativas se fundem, os núcleos haplóides compartilham o mesmo citoplasma em um heterocácion. Nesse exemplo, os núcleos haplóides com mutações em genes diferentes na via de síntese de arginina complementam-se para produzir uma *Neurospora* que não requer mais arginina.

Aqui, entretanto, uma F_2 mostra a proporção padrão 9:3:3:1:

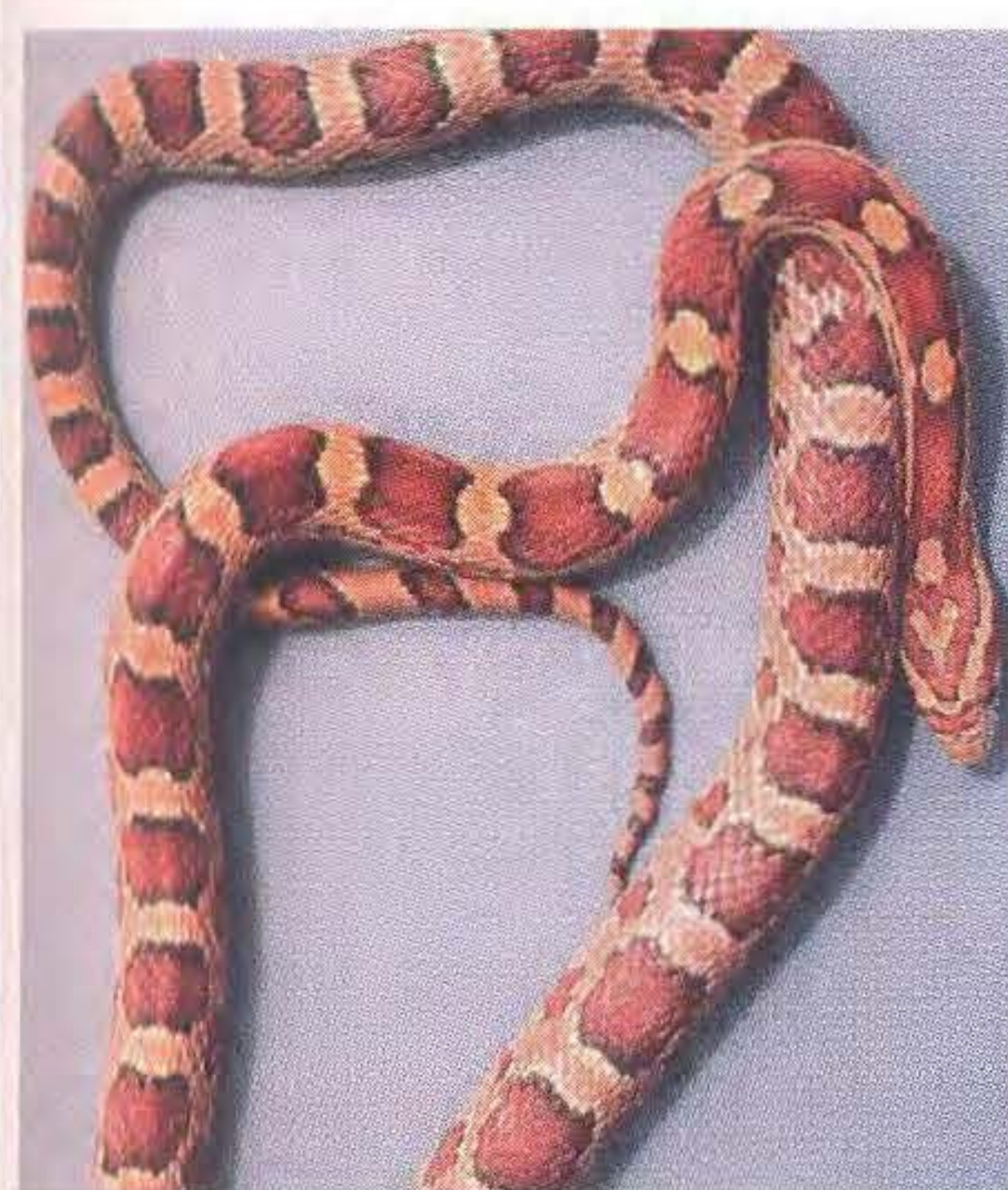
♀ $o^+/o ; b^+/b$ × ♂ $o^+/o ; b^+/b$
(camuflagem) (camuflagem)
↓
 F_2 9 $o^+/- ; b^+/-$ (camuflagem)
3 $o^+/- ; b/b$ (laranja)
3 $o/o ; b^+/-$ (preto)
1 $o/o ; b/b$ (albino)

A proporção 9:3:3:1 é produzida porque os dois genes para pigmento atuam independentemente a nível celular.

precursor $\xrightarrow{b^+}$ pigmento preto
precursor $\xrightarrow{o^+}$ pigmento laranja } camuflagem

Se a presença de um mutante faz uma via falhar, a outra via ainda está ativa, produzindo o outro pigmento de cor. Apenas quando ambos os mutantes estão presentes, ambas as vias falham, e não é produzido pigmento de nenhuma cor.

Pigmentos herdados e sintetizados independentemente



(a)



(b)



(c)



(d)

FIG. 6.17 Nas cobras de milharal, as combinações de pigmentos laranja e preto determinam os quatro fenótipos mostrados. (a) Uma cobra camuflada tipo selvagem preta e laranja produz ambos os pigmentos preto e laranja. (b) Uma cobra preta não produz pigmento laranja. (c) Uma cobra laranja não produz pigmento preto. (d) Uma cobra albina não produz pigmento preto nem laranja. [Anthony Griffiths.]

A proporção 9:7: genes na mesma via A proporção de F_2 do cruzamento da campainha mostra ambas as plantas, azul e branca, em uma proporção de 9:7. Como podem ser explicados tais resultados? A proporção 9:7 é claramente uma modificação da proporção diíbrida 9:3:3:1 com 3:3:1 combinada para fazer 7; assim, é deduzida alguma forma de interação. O cruzamento de duas linhagens brancas e a subsequente geração podem ser representados do seguinte modo:

$w1/w1 ; w2^+/w2^+$ (branca) × $w1^+/w1^+ ; w2/w2$ (branca)

F_1 $w1^+/w1 ; w2^+/w2$ (azul)

$w1^+/w1 ; w2^+/w2$ × $w1^+/w1 ; w2^+/w2$

F_2 9 $w1^+/- ; w2^+/-$ (azul) 9
3 $w1^+/- ; w2/w2$ (branca) } 7
3 $w1/w1 ; w2^+/-$ (branca)
1 $w1/w1 ; w2/w2$ (branca)

Claramente, nesse caso, o único modo no qual uma proporção 9:7 é possível é se o duplo mutante tem os mesmos fenótipos que os mutantes simples. Assim, a proporção modificada constitui um modo de identificar o fenótipo do mutante duplo. Além disso, os fenótipos idênticos dos mutantes único e duplo sugerem que cada alelo mutante controla uma etapa diferente da *mesma* via. Os resultados mostram que uma planta terá pétalas brancas se for homozigota para o alelo mutante recessivo de *um dos* genes ou *ambos* os genes. Para ter o fenótipo azul, uma planta deve ter pelo menos uma cópia do alelo dominante de ambos os genes porque ambos são necessários para completar as etapas sequenciais na via.

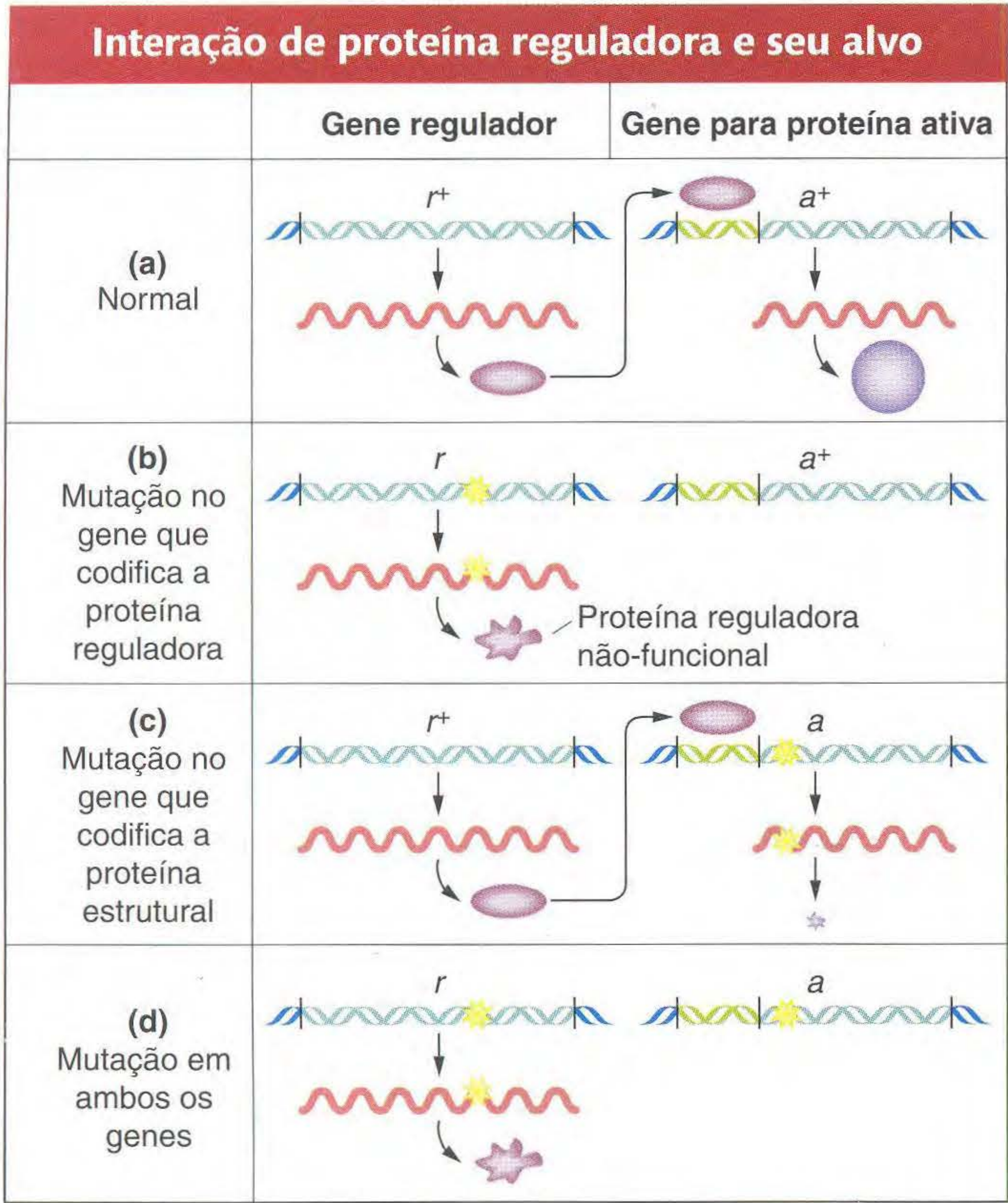


FIG. 6.18 O gene r^+ codifica uma proteína reguladora, e o gene a^+ codifica uma proteína estrutural. Ambos devem ser normais para que seja produzida uma proteína estrutural funcional ("ativa").

Não importa qual seja o ausente, a mesma via falha, produzindo o mesmo fenótipo. Assim, três das classes genótípicas produzirão o mesmo fenótipo, e, assim, de um modo geral, resultam apenas dois fenótipos.

O exemplo em campainha mostra etapas diferentes em uma via sintética. Resultados similares podem surgir da regulação gênica. Um gene regulador em geral funciona produzindo uma proteína que se liga a um sítio regulador de um gene-alvo, facilitando a transcrição do gene (Figura 6.18). Na ausência da proteína reguladora, o gene-alvo seria transcrito em níveis muito baixos, inadequados para as necessidades celulares. Vamos cruzar uma linhagem pura r/r defectiva da proteína reguladora com uma linhagem pura a/a defectiva da proteína-alvo. O cruzamento é $r/r ; a^+/a^+ \times r^+/r^+ ; a/a$. O diíbrido $r^+/r ; a^+/a$ apresentará complementação entre os genótipos mutantes porque tanto r^+ como a^+ estão presentes, permitindo a transcrição normal do alelo tipo selvagem. Quando autofecundados, os diíbridos de F_1 também resultam em uma proporção fenotípica 9:7 na F_2 :

Proporção	Genótipo	Proteína a^+ funcional	Proporção
$\frac{9}{16}$	$r^+/- ; a^+/-$	Sim	9
$\frac{3}{16}$	$r^+/- ; a/a$	Não	7
$\frac{3}{16}$	$r/r ; a^+/-$	Não	
$\frac{1}{16}$	$r/r ; a/a$	Não	

Uma proporção de F_2 de 9:7 sugere interação de genes na mesma via; a ausência de uma das funções leva à ausência do produto final da via.

A proporção 9:3:4: epistasia recessiva Uma proporção 9:3:4 na F_2 sugere um tipo de interação chamada de **epistasia**. Essa palavra significa "ficar sobre", referindo-se à situação na qual o mutante duplo mostra o fenótipo de uma mutação mas não do outro. A mutação que "fica sobre" é **epistática**, enquanto a que "fica sob" é **hipostática**. A epistasia resulta de os genes estarem na mesma via. Em uma via de síntese simples, a mutação epistática é levada por um gene que está mais antecedente (anterior na via) que o gene da mutação epistática. O fenótipo mutante do gene antecedente tem precedência, não importa o que esteja ocorrendo depois na via.

Vejamos um exemplo que envolve a síntese do pigmento da pétala na planta *blue-eyed Mary* (*Collinsia parviflora*). A partir do tipo selvagem azul, começamos com duas linhagens mutantes puras, uma com branca (w/w) e outra com pétalas magenta (m/m). Os genes w e m não estão ligados. A F_1 e F_2 são as seguintes:

$$\begin{array}{l} w/w ; m^+/m^+ \text{ (branca)} \times w^+/w^+ ; m/m \text{ (magenta)} \\ F_1 \quad \quad \quad w^+/w ; m^+/m \text{ (azul)} \\ \downarrow \\ w^+/w ; m^+/m \times w^+/w ; m^+/m \\ \downarrow \\ F_2 \quad \quad \quad \begin{array}{ll} 9 \ w^+/- ; m^+/- & \text{(azul)} & 9 \\ 3 \ w^+/- ; m/m & \text{(magenta)} & 3 \\ 3 \ w/w ; m^+/- & \text{(branca)} & \} \\ 1 \ w/w ; m/m & \text{(branca)} & \} \end{array} \quad 4 \end{array}$$

Na F_2 , a proporção fenotípica 9:3:4 é diagnóstica de epistasia recessiva. Como no caso anterior, vemos, novamente, que a proporção nos diz qual o fenótipo do duplo deve ser, pois o componente 4/16 da proporção deve ser um agrupamento de uma única classe mutante (3/16) mais a classe mutante dupla (1/16). Assim, o mutante duplo expressa apenas um dos dois fenótipos mutantes; logo, por definição, branco deve ser epistático em relação a magenta. (Para encontrar o mutante duplo dentro do grupo 4/16, as plantas brancas F_2 teriam que ser individualmente submetidas a cruzamento-teste.) Essa interação é chamada de epistasia recessiva porque o fenótipo recessivo (branco) supera o outro fenótipo. A epistasia dominante será considerada na seção seguinte.

A nível celular, podemos considerar a epistasia recessiva em *Collinsia* pelo seguinte tipo de via (veja também Figura 6.19).



Note que a mutação epistática ocorre em uma etapa da via que leva a um pigmento azul; essa etapa é antecedente à etapa que está bloqueada por mutação oculta.

Um modelo para epistasia recessiva

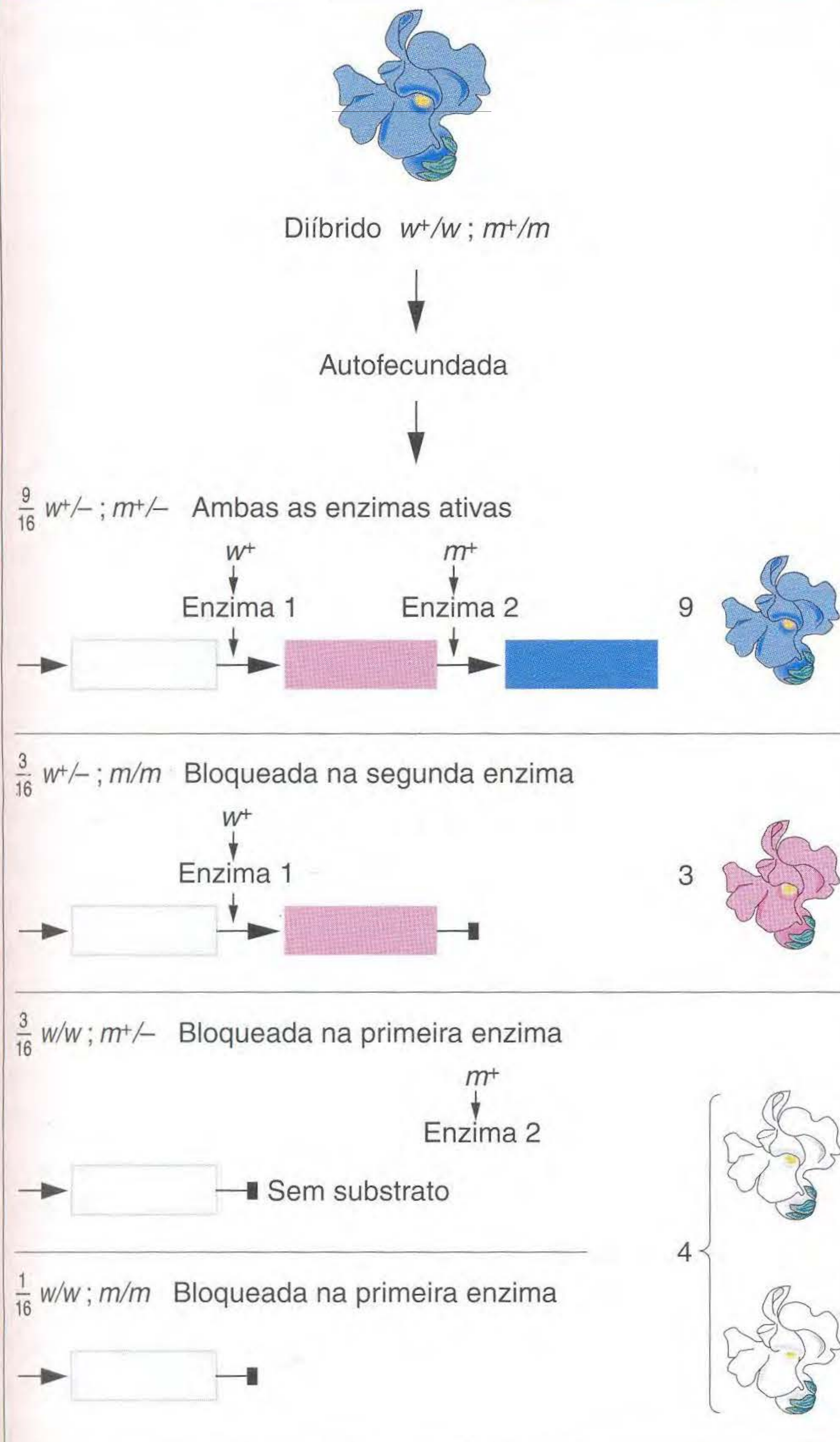


FIG. 6.19 Alelos tipo selvagem de dois genes (w^+ e m^+) codificam enzimas que catalisam etapas sucessivas na síntese de um pigmento azul de pétala. As plantas homozigotas m/m produzem flores magenta, e as plantas homozigotas w/w produzem flores brancas. O mutante duplo $w/w; m/m$ também produz flores brancas, indicando que branca é epistática em relação a magenta.

Outro caso informativo de epistasia recessiva é a cor de pelagem amarela de alguns cães da raça Labrador. Dois alelos, B e b , para pelagens preta e marrom, respectivamente. Os dois alelos produzem melanina preta e marrom. O alelo e de outro gene é epistático em relação a esses alelos, dando uma pelagem amarela (Figura 6.20). Portanto, os genótipos $B/-; e/e$ e $b/b; e/e$ produzem um fenótipo amarelo, enquanto $B/-; E/-$ e $b/b; E/-$ são preto e marrom, respectivamente. Esse caso de epistasia não é causado por um bloqueio antecedente em uma via que leva ao pigmento escuro. Os cães amarelos podem fazer pigmento marrom, como pode ser visto no nariz e lábios. A ação do alelo e é evitar a deposição do pigmento nos pêlos. Nesse caso, o gene epistático está *desenvolvimentalmente posterior*; ele representa um tipo de alvo desenvolvimental que deve ser do genótipo E antes que o pigmento seja depositado.

A epistasia é deduzida quando um alelo mutante de um gene mascara a expressão de um alelo mutante de outro gene e expressa seu próprio fenótipo.

Em fungos, a análise de tétrades é útil na identificação de um mutante duplo. Por exemplo, um asco contendo metade de seus produtos como tipo selvagem deve conter mutantes duplos. Considere o cruzamento

$$a \cdot b^+ \times a^+ \cdot b$$

Em certa proporção da prole, os alelos a e b vão segregar juntos (um asco ditipo não-parental). Tal tétrade apresentará os seguintes fenótipos:

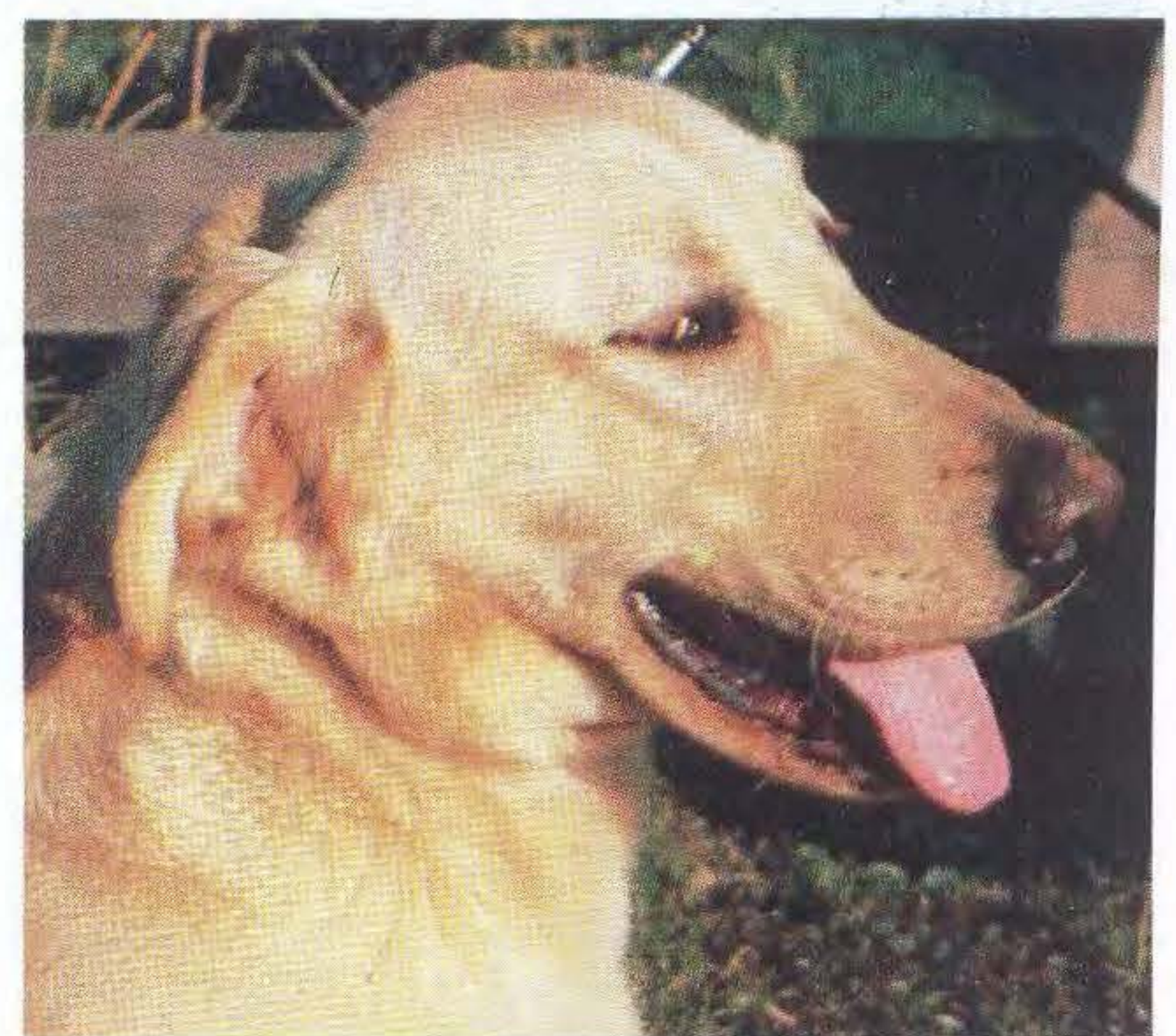
Epistasia recessiva devida à mutação de pelagem amarela



(a)



(b)



(c)

FIG. 6.20 Três diferentes cores de pelagem em labrador. Dois alelos B e b de um gene de pigmento determinam (a) preto e (b) marrom, respectivamente. Em um gene separado, E permite a deposição de cor na pelagem, e e/e impede a deposição, resultando em (c) o fenótipo dourado. A parte c ilustra a epistasia recessiva. [Anthony Griffiths.]

tipo selvagem	$a^+ \cdot b^+$	mutante duplo	$a \cdot b$
tipo selvagem	$a^+ \cdot b^+$	mutante duplo	$a \cdot b$

Assim, o mutante duplo deve ter o genótipo não-tipo selvagem e pode ser avaliado de acordo. Se o fenótipo é a , então b está sendo superado; se o fenótipo é b , então a está sendo superado. Se ambos os fenótipos estão presentes, então não há epistasia.

A proporção 12:3:1: epistasia dominante Em dedaleira (*Digitalis purpurea*), dois genes interagem na via que determina a cor da pétala. Os dois genes não são ligados. Um gene afeta a intensidade do pigmento vermelho na pétala; o alelo d resulta na cor vermelho-clara visto nas populações naturais de dedaleira, enquanto D é um alelo mutante que produz a cor vermelho-escura (Figura 6.21). O outro gene determina em quais células o pigmento é sintetizado: o alelo w permite a síntese do pigmento nas pétalas como no tipo selvagem, mas o alelo mutante W confina a síntese do pigmento a pequenos pontos. Se autofecundarmos um diíbrido $D/d ; W/w$, então a proporção de F_2 será a seguinte:

9 $D/- ; W/-$ branca com pontos	}	12
3 $d/d ; W/-$ branca com pontos		
3 $D/- ; w/w$ vermelho-escura		3
1 $d/d ; w/w$ vermelho-clara		1

A proporção nos diz que o alelo dominante W é epistático, produzindo a proporção 12:3:1. O componente 12/16 da proporção deve incluir a classe mutante dupla (9/16), que é claramente branca em fenótipo, estabelecendo a epistasia do alelo dominante W . Os dois genes atuam em uma via desenvolvimental comum: W impede a síntese do pigmento vermelho, mas apenas em uma classe especial de células que constituem a área principal da pétala; a síntese é permitida nos pontos da garganta. Quando a síntese é permitida, o pigmento pode ser produzido em concentrações altas ou baixas.



FIG. 6.21 Em dedaleira, D e d causam pigmentos escuros e claros, respectivamente, enquanto o epistático W restringe o pigmento a pontos na garganta. [Anthony Griffiths.]

Supressores Não é fácil selecionar especificamente ou triar interações epistáticas, e os casos de epistasia têm que ser construídos pela combinação laboriosa de mutações candidatas duas de cada vez. Entretanto, para nosso tipo seguinte de interação gênica, o experimentador pode selecionar prontamente alelos mutantes interessantes. Um **supressor** é um alelo mutante de um gene que reverte o efeito de uma mutação de outro gene, resultando em um fenótipo tipo selvagem ou quase tipo selvagem. A supressão implica que os produtos gênicos normalmente interagem. Por exemplo, suponha que um alelo a^+ produza o fenótipo normal, enquanto um alelo mutante recessivo a resulta em anormalidade. Um alelo mutante recessivo s em outro gene suprime o efeito de a , e, logo, o genótipo $a/a \cdot s/s$ terá o fenótipo tipo selvagem (tipo a^+). Os alelos supressores às vezes não têm efeito na ausência de outra mutação; em tal caso, o fenótipo de $a^+/a^+ \cdot s/s$ seria tipo selvagem. Em outros casos, o alelo supressor produz seu próprio fenótipo anormal.

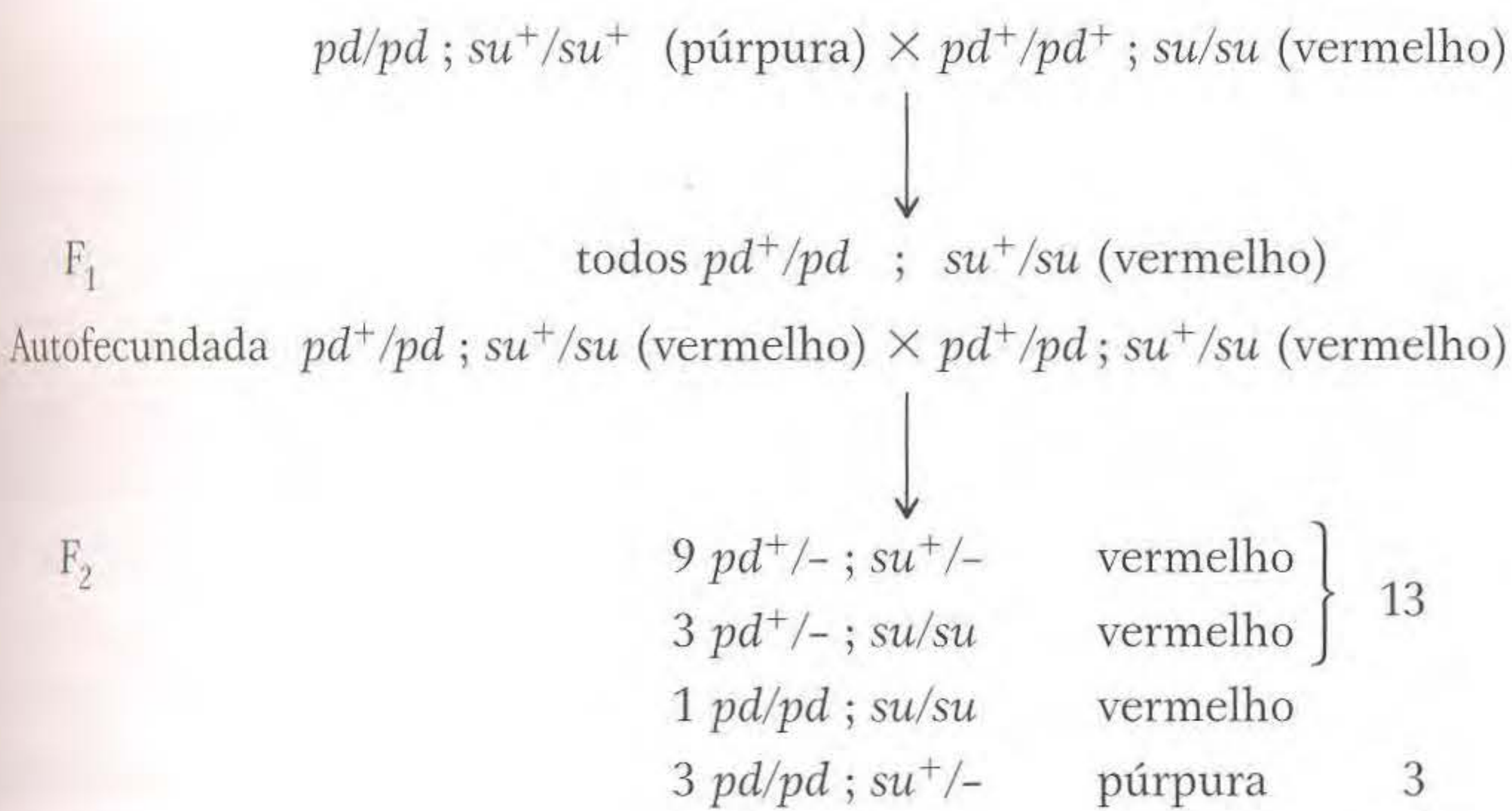
A triagem de supressores é bem direta. Comece com um mutante em algum processo de interesse, exponha esse mutante a agentes causadores de mutação tal como radiação de alta energia, e faça uma triagem dos descendentes para os tipos selvagens. Em haplóides tais como fungos, a triagem é feita simplesmente plaqueando células mutagenizadas e procurando colônias com fenótipos tipo selvagem. Os tipos mais selvagens que surgem desse modo são meras reversões do evento mutacional original, e são chamados de **revertentes**. Entretanto, alguns serão “pseudo-revertentes”, mutantes duplos nos quais uma das mutações é um supressor.

Revertentes e estados suprimidos podem ser distintos por cruzamento apropriado. Por exemplo, em leveduras, os dois resultados seriam distintos do seguinte modo:

revertente verdadeiro $a^+ \times$ tipo selvagem padrão a^+		
	↓	
Prole	todos a^+	
mutante suprimido $a \cdot s \times$ tipo selvagem padrão $a^+ \cdot s^+$		
	↓	
Prole	$a^+ \cdot s^+$	tipo selvagem
	$a^+ \cdot s$	tipo selvagem
	$a \cdot s^+$	mutante original
	$a \cdot s$	tipo selvagem (suprimido)

O surgimento do fenótipo mutante original identifica o genitor como mutante suprimido.

Em diplóides, os supressores produzem proporções F_2 modificadas, que são úteis para confirmar a supressão. Vejamos um exemplo da vida real de *Drosophila*. O alelo recessivo pd resulta em olhos púrpura quando não-suprimido. Um alelo recessivo su não tem em si um fenótipo detectável, mas suprime o alelo recessivo não-ligado pd . Assim, $pd/pd ; su/su$ tem aspecto tipo selvagem e olhos vermelhos. A seguinte análise ilustra o padrão de herança. Uma mosca homozigota para olhos púrpura é cruzada com uma homozigota de olhos vermelhos levando o supressor.

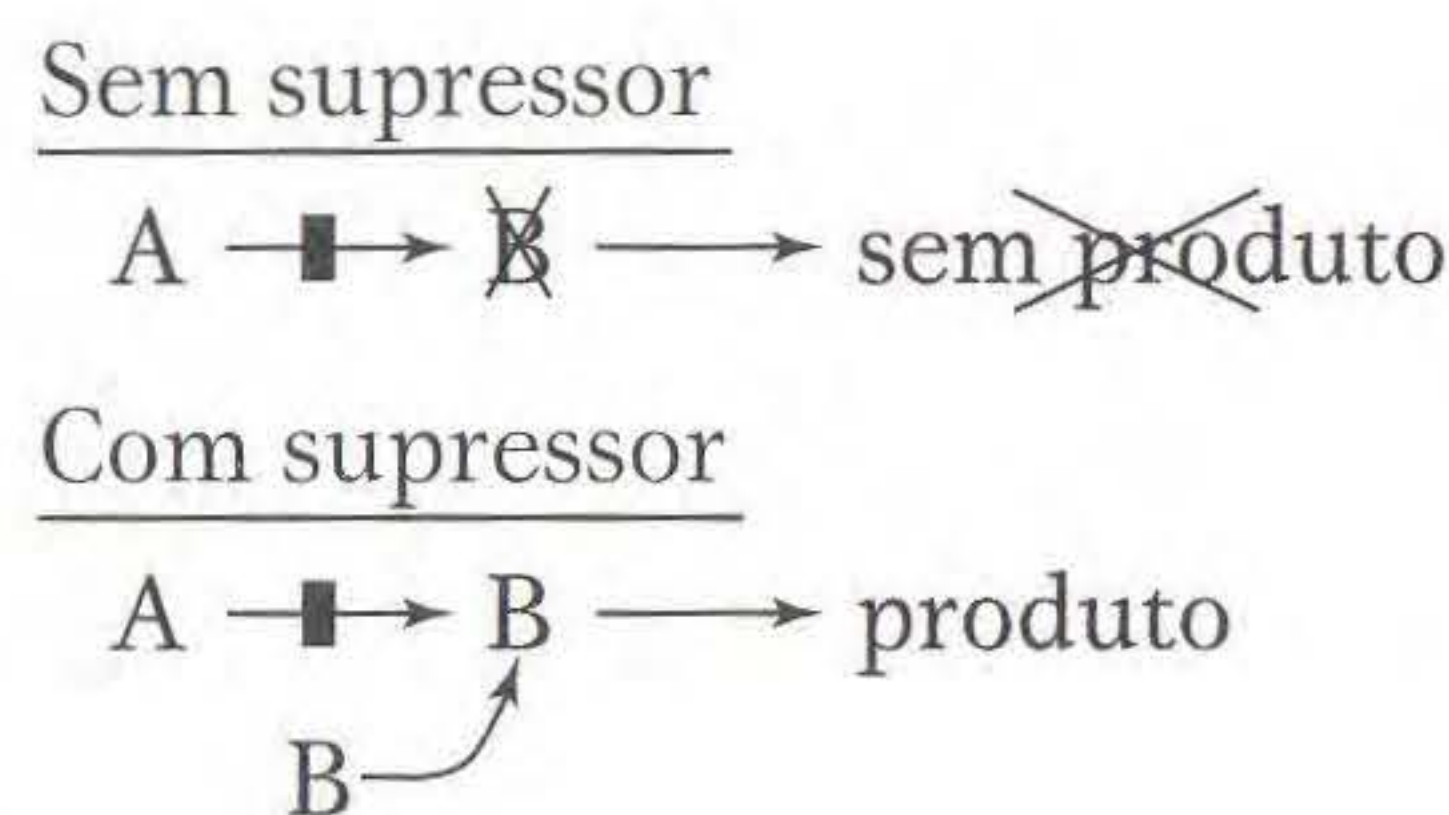


A proporção geral na F_2 é 13 vermelhos : 3 púrpura. O componente 13/16 deve induzir o mutante duplo, que é claramente de fenótipo tipo selvagem. Essa proporção é característica de um supressor recessivo agindo em uma mutação recessiva. São encontrados tanto supressores recessivos quanto dominantes, e eles podem agir em mutações recessivas ou dominantes. Essas possibilidades resultam em uma variedade de proporções fenotípicas diferentes.

A supressão é às vezes confundida com epistasia. Entretanto, a diferença importante é que um supressor cancela a expressão de um alelo mutante e restaura o fenótipo tipo selvagem correspondente. Além disso, em geral apenas dois fenótipos segregam (como nos exemplos anteriores) em vez de três, como na epistasia.

Como os supressores funcionam a nível molecular? Existem muitos mecanismos possíveis. Um tipo particularmente útil de supressão é baseado na ligação física dos produtos gênicos na célula — por exemplo, ligação proteína-proteína. Suponha que duas proteínas normalmente se ajustem para dar algum tipo de função celular. Quando uma mutação causa uma mudança de forma em uma proteína, ela não se ajusta mais à outra; assim, a função é perdida (Figura 6.22). Entretanto, uma mutação supressora que causa uma mudança compensatória na segunda proteína pode restaurar o ajuste e, portanto, o funcionamento normal. Desse modo podem ser deduzidas proteínas que interagem.

Alternativamente, em situações nas quais uma mutação causa um bloqueio em uma via metabólica, o supressor encontra um meio de ultrapassar o bloqueio — por exemplo, redirecionando a via bloqueada para a que está além do bloqueio. No exemplo seguinte, o supressor fornece um intermediário B para evitar o bloqueio.



Em vários organismos, foram encontrados *supressores em sentido*, mutações em genes de tRNA que resultam em um anticódon que irá ligar-se a um códon de fim prematu-

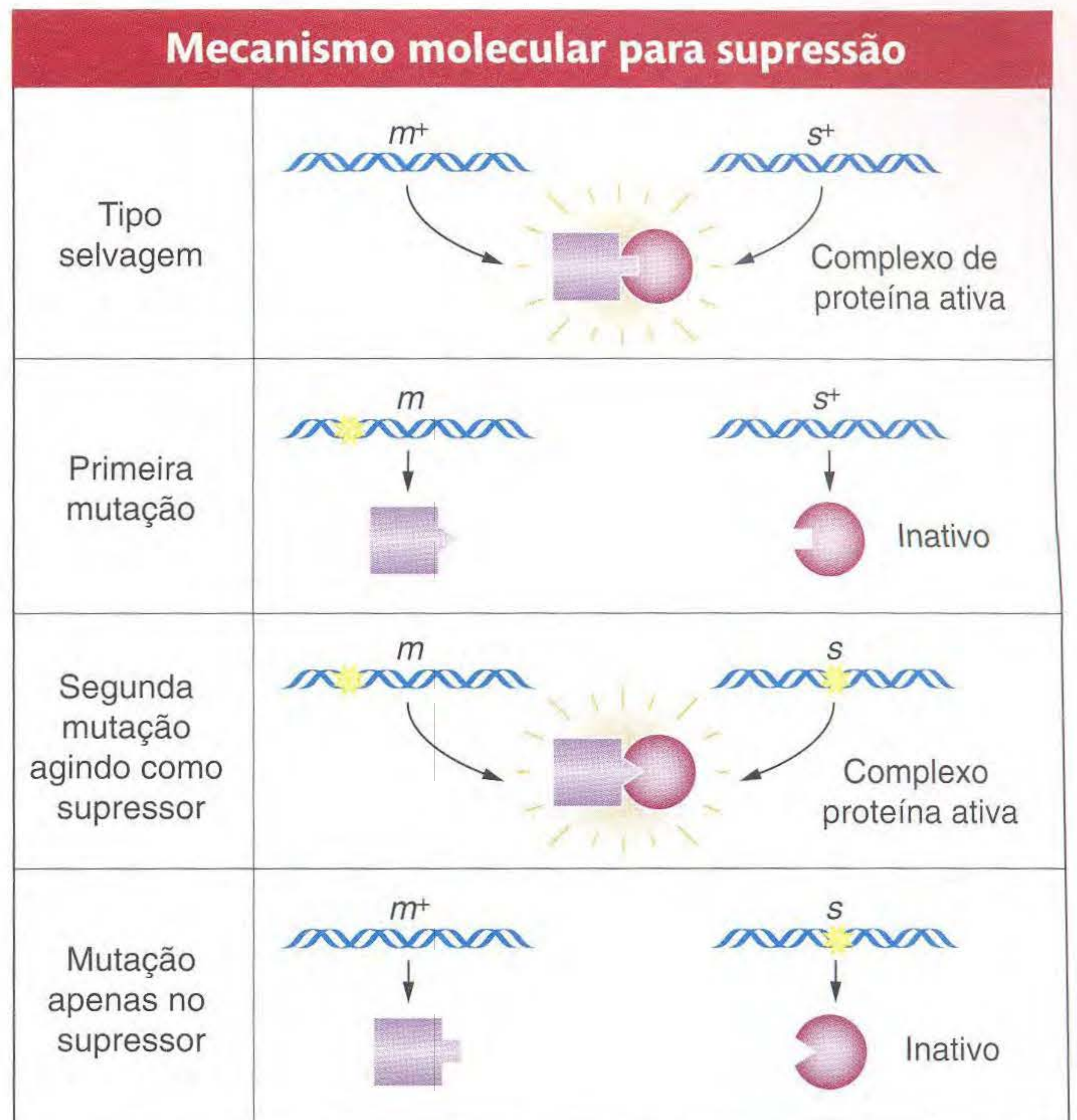


FIG. 6.22 Uma primeira mutação altera o sítio de ligação de uma proteína, de modo que ela não pode mais se ligar a um parceiro. Uma mutação supressora no parceiro altera o sítio de ligação, de modo que as proteínas são capazes de se ligar novamente.

ro dentro de uma sequência codificante mutante. Assim, o supressor permite que a tradução continue além do bloqueio anterior.

Alelos mutantes chamados de supressores cancelam a expressão de um alelo mutante de outro gene, resultando no fenótipo tipo selvagem normal.

Modificadores Como o nome sugere, uma mutação **modificadora** em um segundo locus muda o grau de expressão de um gene mutado no primeiro locus. Os genes reguladores fornecem uma ilustração simples. Como em um exemplo anterior, as proteínas reguladoras ligam-se ao DNA antecedente ao ponto de início da transcrição. Essas proteínas regulam o nível da transcrição. Na discussão da complementação, consideramos uma mutação nula de um gene regulador que evita quase que completamente a transcrição. Entretanto, algumas mutações reguladoras mudam o nível da transcrição do gene-alvo, de modo que é produzida mais ou menos proteína. Em outras palavras, uma mutação em uma proteína reguladora pode diminuir ou aumentar o transcrito do gene. Vejamos um exemplo usando uma mutação *b* que regula para menos, afetando o gene *A* em um fungo tal como uma levedura. Procuramos o efeito de *b* em uma mutação *leaky* (vazadora) no gene *A*, que é uma mutação que produz uma proteína com alguma baixa função residual. Cruzamos uma mutação *leaky a* com a mutação reguladora *b*:

mutante *leaky* $a^+ \cdot b^+ \times$ regulador ineficiente $a^+ \cdot b$

Prole	Fenótipo
$a^+ \cdot b^+$	tipo selvagem
$a^+ \cdot b$	defectivo (baixa transcrição)
$a \cdot b^+$	defectivo (proteína A defectiva)
$a \cdot b$	extremamente defectiva (baixa transcrição da proteína defectiva)

Assim, a ação do modificador é vista no surgimento de duas gradações de fenótipos mutantes *dentro* da prole *a*.

Letais sintéticos Em alguns casos, quando dois mutantes únicos viáveis são entrecruzados, os mutantes duplos resultantes são letais. Em um diplóide de F₂, esse resultado seria manifesto como uma proporção 9:3:3 porque o mutante duplo (que seria o componente “1” da proporção) estaria ausente. Esses **letais sintéticos** podem ser considerados uma categoria especial de interação gênica. Eles podem indicar tipos especiais de interações de produtos gênicos. Por exemplo, a análise do genoma revelou que a evolução produziu muitos sistemas duplicados dentro da célula. Uma vantagem dessas duplicações pode ser fornecer “reservas” (*backups*). Se existem mutações nulas em genes em ambos os sistemas duplicados, então um sistema defeituoso não teria reserva, e o indivíduo não teria uma função essencial e morreria. Em outro caso, uma mutação *leaky* em uma etapa de uma via pode fazer com que

a via diminua, mas deixe uma função suficiente para viver. Entretanto, se os mutantes duplos se combinam, cada um com uma mutação *leaky* em uma etapa diferente, toda a via para-ria. Uma versão desta última interação é de duas mutações em uma máquina de proteína, como mostrado na Figura 6.23.

A construção de uma máquina de proteína é, em parte, uma questão de proteínas constituintes encontrando-se por movimento molecular aleatório e ligando-se por formas complementares. Entretanto, a montagem de algumas máquinas de proteína é como uma pequena via com etapas que requerem energia e enzimas. Tal exemplo é mostrado na Figura 6.24. Qualquer um desses componentes que intera-gem, sejam componentes da máquina ou enzimas associadas, pode ser dissecado pela análise de letais sintéticos.

Na discussão anterior de proporções mendelianas modifica-das, todos os cruzamentos foram autofecundações de diíbridos.

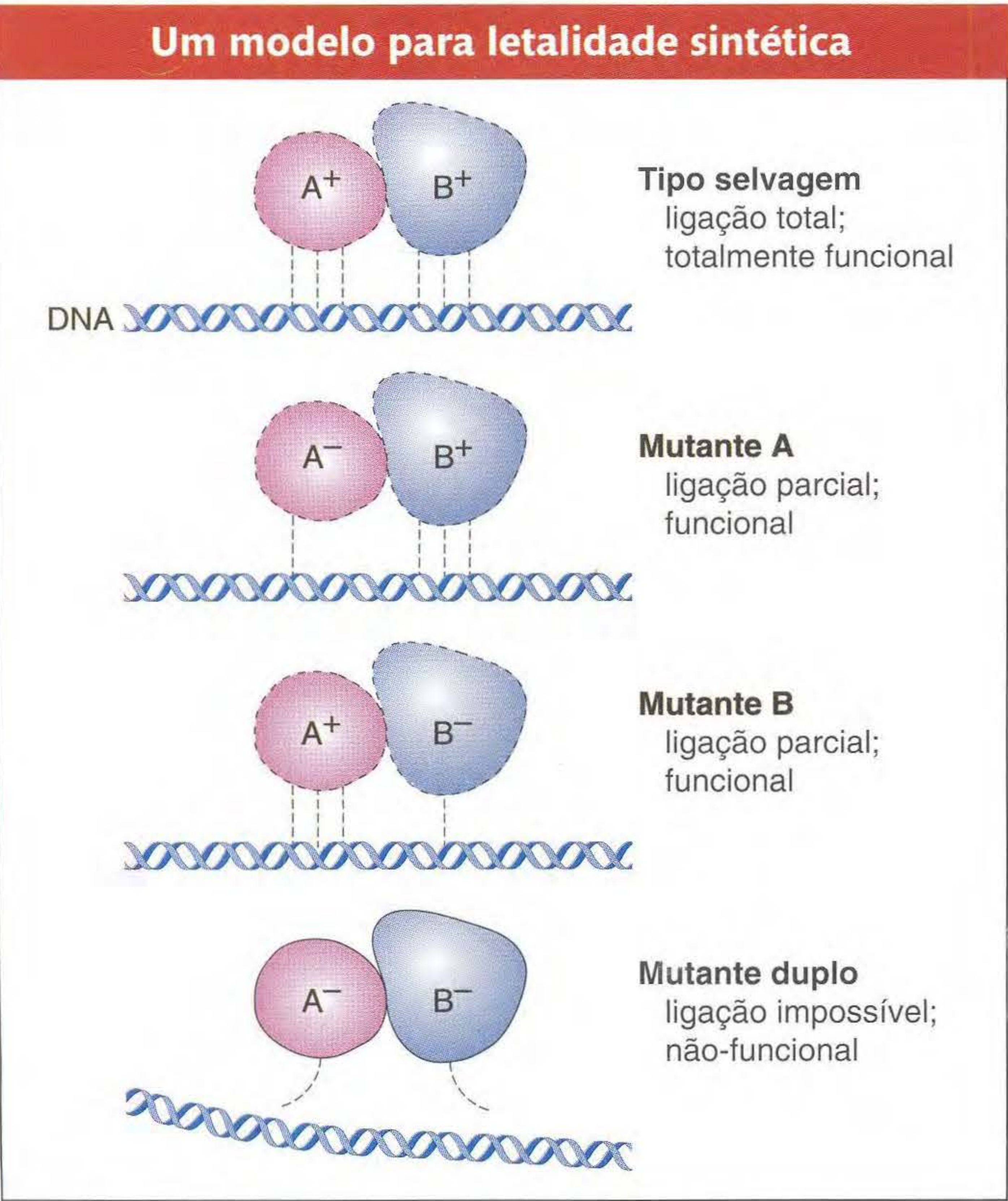


FIG. 6.23 Duas proteínas interagindo desempenham alguma função essencial em algum substrato, tal como o DNA, mas primeiro deve se ligar a ele. A ligação reduzida de uma das proteínas permite que reste alguma funcionalidade, mas a ligação reduzida de ambas é letal.

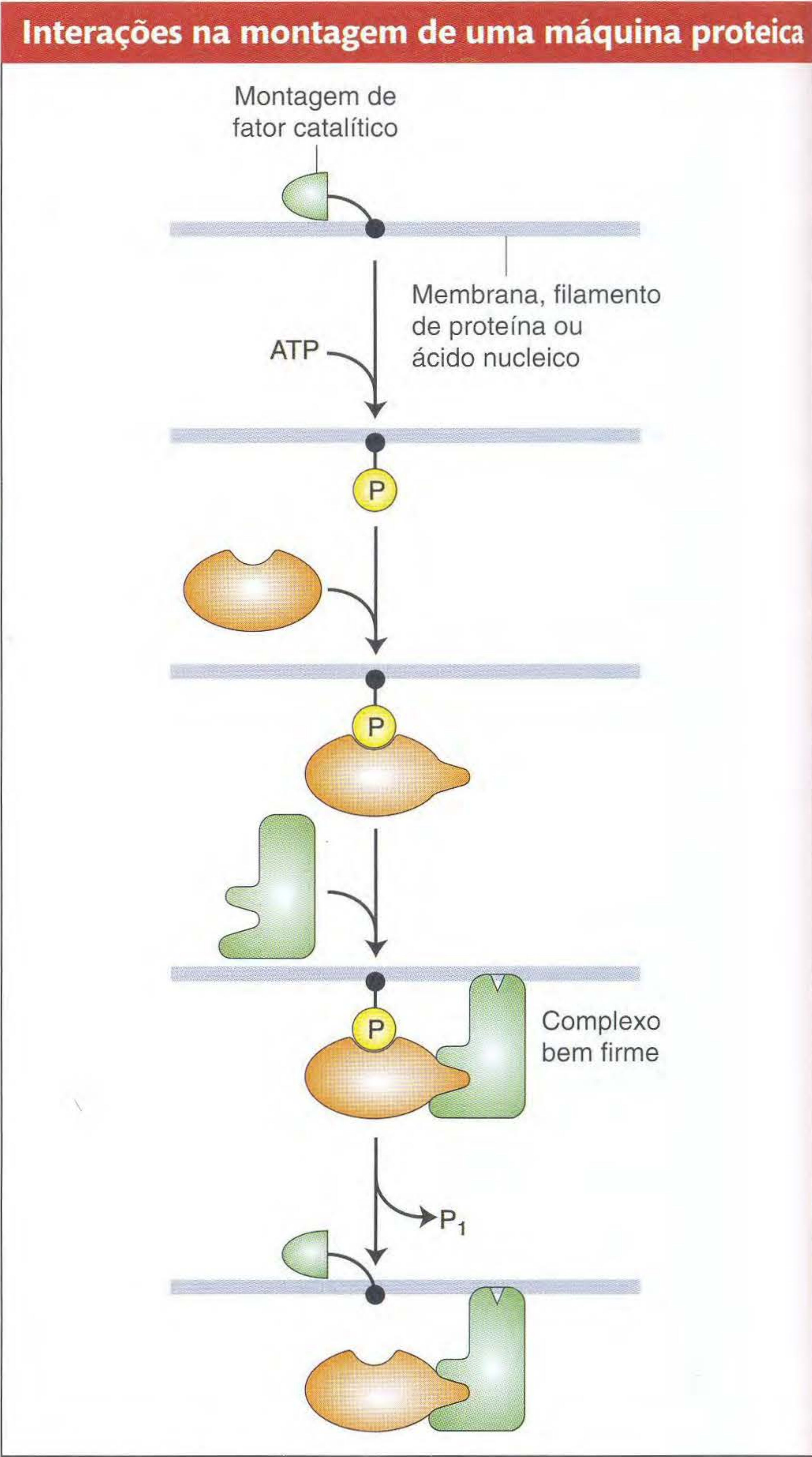


FIG. 6.24 A fosforilação ativa um fator de montagem, permitindo que as máquinas proteicas sejam montadas *in situ* em uma membrana, fila-mento ou ácido nucleico. [De B. Alberts, “The Cell As a Collection of Protein Machines,” *Cell* 92, 1998, 291–294.]

Como um exercício, você pode querer calcular as proporções que seriam produzidas nos mesmos sistemas se fossem feitos cruzamentos-teste em vez de autofecundações.

6.4 Penetrância e Expressividade

Na análise da herança monogênica, há uma tendência natural para escolher mutantes que produzem claras proporções mendelianas. Em tais casos, podemos usar o fenótipo para distinguir genótipos mutantes e tipo selvagem com quase 100% de certeza. Nesses casos, dizemos que a mutação é 100% *penetrante* no fenótipo. Entretanto, muitas mutações mostram penetrância *incompleta*; isto é, nem todos os indivíduos com o genótipo expressam o fenótipo correspondente. Assim, a **penetrância** é definida como a percentagem de *indivíduos* com determinado alelo que exibem o fenótipo associado a esse alelo.

Por que um organismo teria determinado genótipo e não expressaria o fenótipo correspondente? Existem vários motivos possíveis:

1. *A influência do meio.* Como dito no Capítulo 1, os indivíduos com o mesmo genótipo podem apresentar uma gama de fenótipos, dependendo do ambiente. A gama de fenótipos para indivíduos mutantes e tipo selvagem podem se superpor: o fenótipo de um indivíduo mutante criado em um conjunto de circunstâncias pode corresponder ao fenótipo de um indivíduo tipo selvagem criado em um conjunto diferente de circunstâncias. Se essa correspondência ocorrer, o mutante pode ser distinto do tipo selvagem.
2. *A influência de outros genes interagentes.* Os modificadores não-caracterizados, genes epistáticos ou supressores no resto do genoma podem agir para evitar a expressão do fenótipo típico.
3. *A sutileza do fenótipo mutante.* Os efeitos sutis causados pela ausência de uma função gênica podem ser difíceis de medir em situação de laboratório.

Um encontro típico com a penetrância incompleta é mostrado na Figura 6.25. Nesse heredograma humano, vemos um fenótipo normalmente herdado de modo dominante desaparecendo na segunda geração apenas para reaparecer na seguinte.

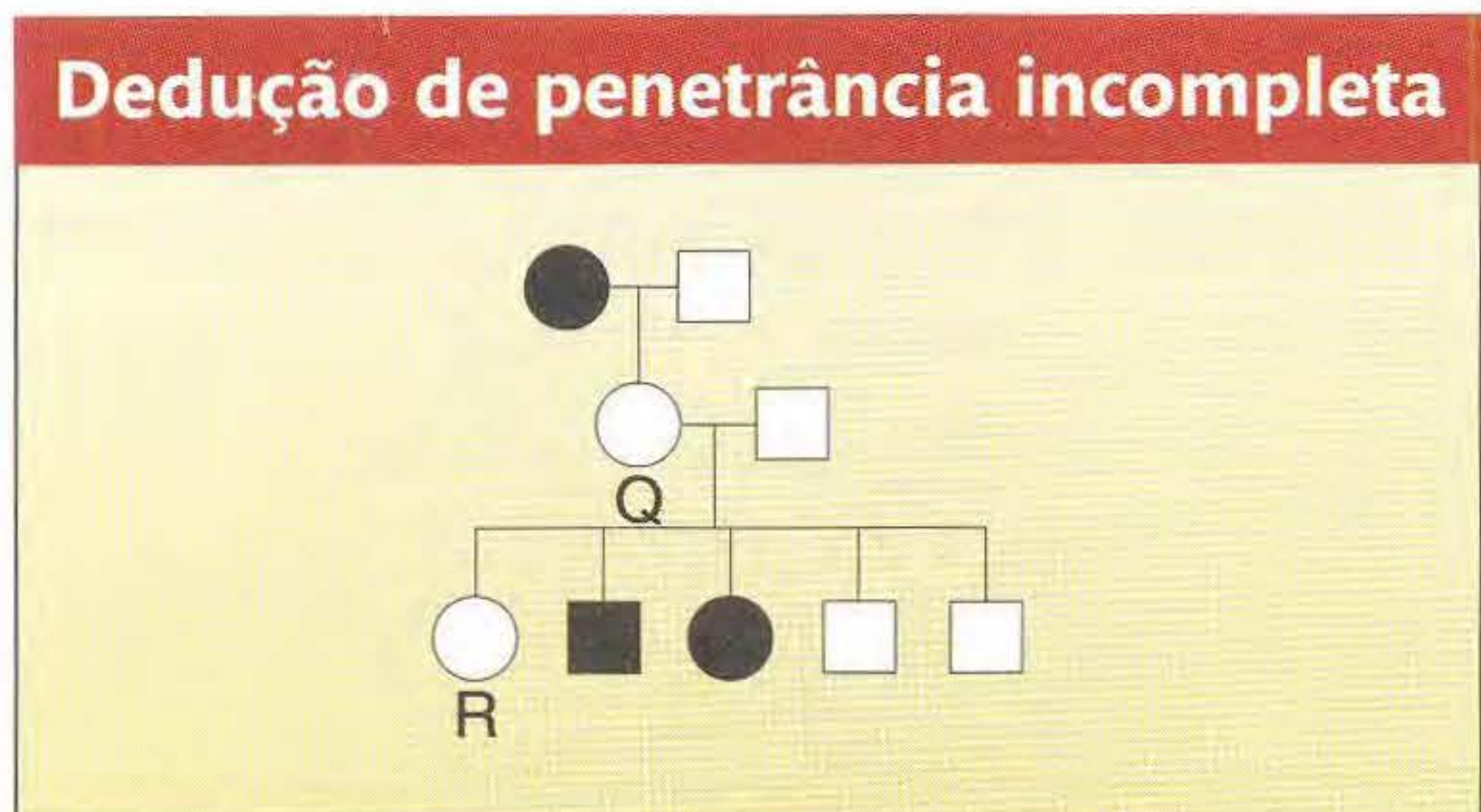


FIG. 6.25 Nesse heredograma de um alelo dominante que não é totalmente penetrante, a pessoa Q não apresenta o fenótipo mas transmite o alelo dominante para, pelo menos, duas da prole. Como o alelo não é totalmente penetrante, a outra prole (por exemplo, R) pode ou não ter herdado o alelo dominante.

Outra medida para descrever a gama de expressão fenotípica é chamada de **expressividade**. A expressividade mede o grau ao qual determinado alelo é expresso a nível fenotípico; isto é, a expressividade mede a intensidade do fenótipo. Por exemplo, os animais “brown” (genótipo b/b) de estoques diferentes podem mostrar intensidades muito diferentes do pigmento marrom (*brown*) (genótipo b/b) de claro a escuro. Como na penetrância, a expressividade variável pode ser devida à variação na constituição alélica do resto do genoma ou a fatores ambientais. A Figura 6.26 ilustra a distinção entre penetrância e expressividade. Um exemplo de expressividade variável em cães é encontrado na Figura 6.27.

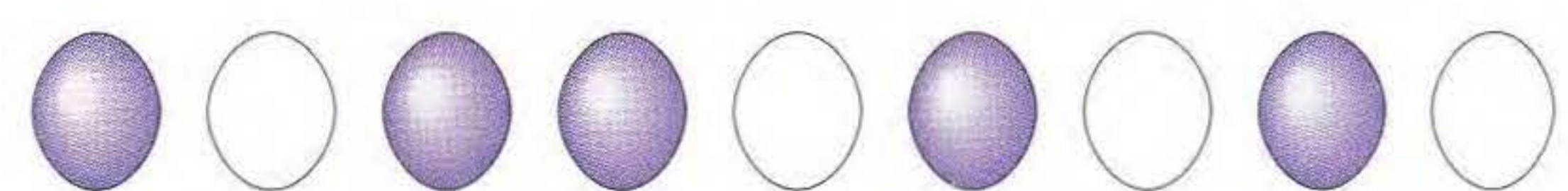
O fenômeno de penetrância incompleta e expressividade variável pode tornar qualquer tipo de análise genética muito mais difícil, incluindo a análise de heredogramas humanos e as previsões na consulta genética. Por exemplo, geralmente é o caso em que um alelo causador de doença não seja totalmente penetrante. Assim, alguém pode ter o alelo mas não apresentar nenhum sinal da doença. Se for o caso, é difícil dar um claro atestado de saúde a qualquer pessoa em um heredograma de doença (por exemplo, a pessoa R na Figura 6.25). Por outro lado, a análise de heredogramas pode, às vezes, identificar pessoas que não expressam mas quase certamente têm um genótipo de doença (por exemplo, o indivíduo Q na Figura 6.25). Similarmente, a expressividade variável pode complicar a consulta porque as pessoas com baixa expressividade podem não ser diagnosticadas.

Muito embora a penetrância e a expressividade possam ser quantificadas, elas representam situações obscuras porque raramente é possível identificar os fatores específicos causando variação sem uma substancial pesquisa extra.

A penetrância e a expressividade quantificam a modificação da expressão gênica variando o ambiente e o fundo genético; elas medem, respectivamente, a percentagem de casos nos quais o gene é expresso e o nível de expressão.

Penetrância e expressividade contrastadas

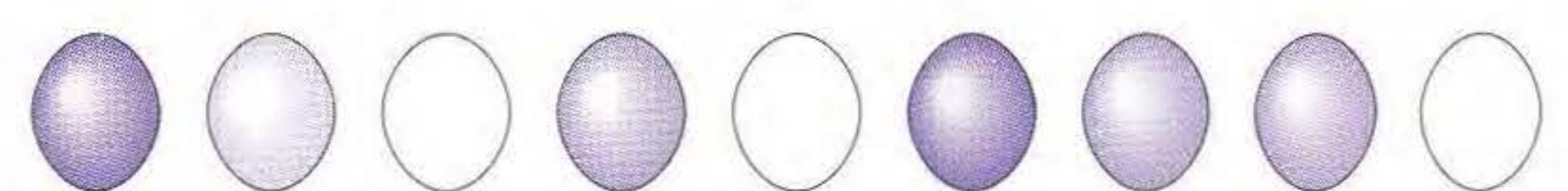
Expressão fenotípica
(cada oval representa um indivíduo)



Penetrância variável



Expressividade variável



Penetrância e expressividade variáveis

FIG. 6.26 Suponha que todos os indivíduos mostrados têm o mesmo alelo de pigmento (P) e possuem o mesmo potencial de produzir o pigmento. Os efeitos do resto do genoma e do ambiente podem suprimir ou modificar a produção de pigmento em qualquer indivíduo. A cor indica o nível de expressão.

Expressividade variável

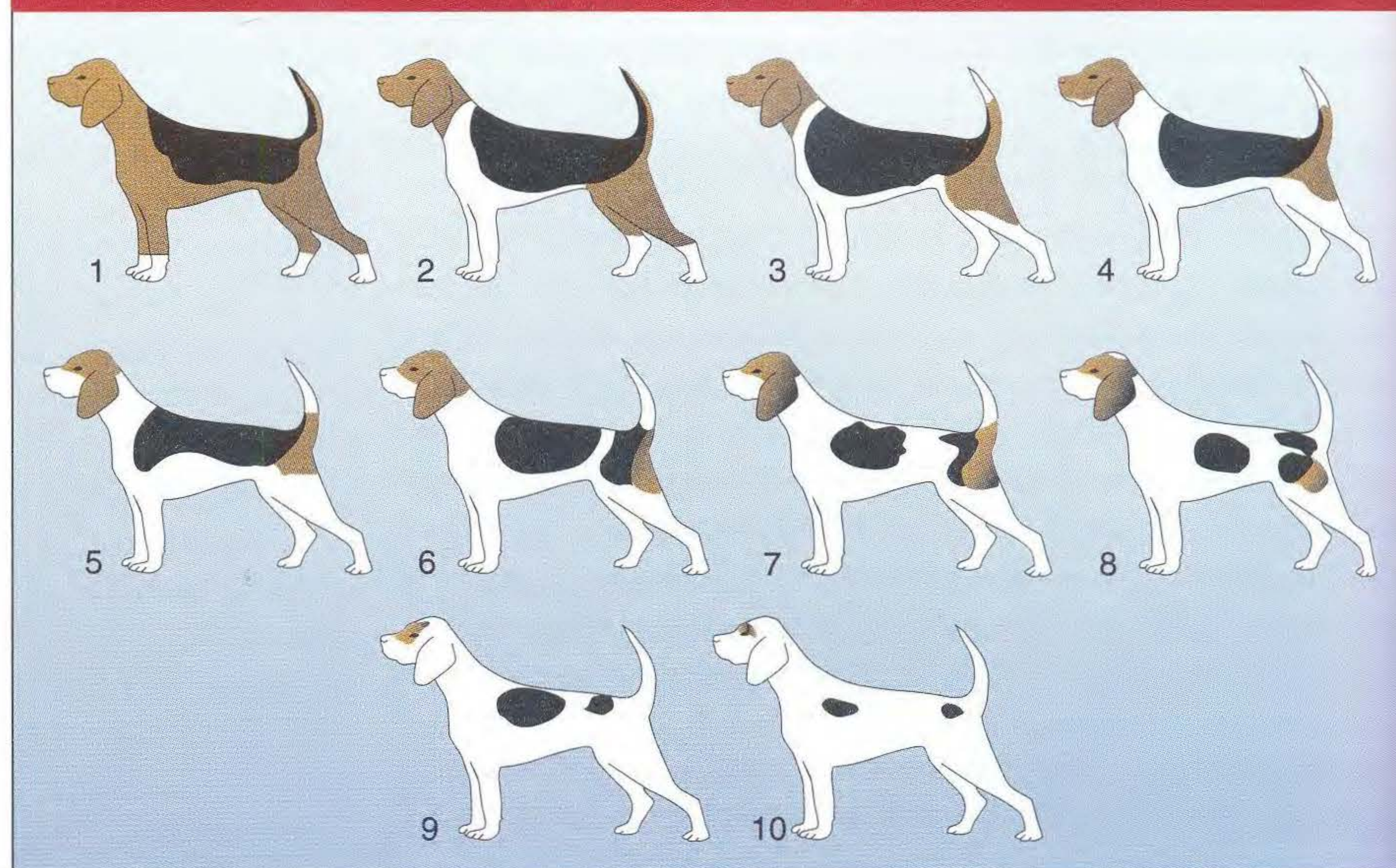


FIG. 6.27 Dez graus de padrões malhados em beagles. Cada um desses cães tem o alelo S^P , o alelo responsável pelas manchas nos cães. A variação é causada pela variação em outros loci. [De Clarence C. Little, *The Inheritance of Coat Color in Dogs*. Cornell University Press, 1957; e Giorgio Schreiber, *J. Hered.* 9, 1930, 403.]

► Resumo

Um gene não age sozinho; ele atua em conjunto com muitos outros genes no genoma. Na análise genética direta, a dedução dessas interações complexas é um estágio importante da pesquisa. As mutações individuais são primeiro testadas quanto a suas relações de dominância, um tipo de interação alélica. As mutações recessivas em geral são o resultado de haplossuficiência do alelo tipo selvagem, enquanto as mutações dominantes são, em geral, o resultado de haploinsuficiência do tipo selvagem ou do mutante agindo como dominante negativo (um polipeptídeo diferente). Algumas mutações causam graves efeitos ou mesmo a morte (mutações letais). A letalidade de uma mutação homozigota recessiva é um modo de avaliar se um gene é essencial no genoma.

A interação de genes diferentes é um resultado de sua participação na mesma via, ou em vias conectantes de vários tipos, de síntese, sinal de transdução ou desenvolvimental. A dissecção genética das interações gênicas começa com o experimentador reunindo mutantes que afetam uma característica de interesse. O teste de complementação determina se duas mutações recessivas distintas são de um gene ou dois genes diferentes. Os genótipos mutantes são juntados em um indivíduo F_1 , e, se o fenótipo é mutante, então não ocorreu complementação e os dois alelos devem estar no mesmo gene. Se o fenótipo é do tipo selvagem,

então a complementação ocorreu e os alelos devem estar em genes diferentes.

A interação de genes diferentes pode ser detectada testando mutantes duplos, pois a interação alélica significa a interação de produtos gênicos a nível funcional. Alguns tipos importantes de interação são a epistasia, supressão e letalidade sintética. A epistasia é a substituição de um fenótipo mutante, produzido por uma mutação por um fenótipo mutante, por uma mutação em outro gene. A observação de epistasia sugere uma via desenvolvimental comum ou química. Um supressor é uma mutação do gene que pode restaurar o fenótipo tipo selvagem a uma mutação em outro gene. Os supressores em geral revelam proteínas de interação física ou ácidos nucleicos. Algumas combinações de mutantes viáveis são letais, um resultado conhecido como *letalidade sintética*. Os letais sintéticos podem revelar uma variedade de interações, dependendo da natureza das mutações.

Os tipos diferentes de interações gênicas produzem proporções diíbridas de F_2 que são modificações do padrão 9:3:3:1. Por exemplo, a epistasia recessiva resulta em uma proporção 9:3:4.

Em termos mais gerais, a interação gênica e a interação ambiental são reveladas pela penetrância variável (a habilidade de um genótipo em se expressar) e a expressividade (o grau quantitativo de expressão de um genótipo).

► Termos-chave

alelo letal
alelo pleiotrópico
alelos múltiplos (séries alélicas)
co-dominância

dominância incompleta
dominância total (completa)
epistasia
expressividade

heterocárion
hipótese um gene - um polipeptídeo
letal sintético
modificador

- mutação dominante negativa

mutação nula

mutações sensíveis à temperatura (ts)

mutantes duplos
- penetrância

revertente

RNA funcional

séries alélicas (vários alelos)
- supressor

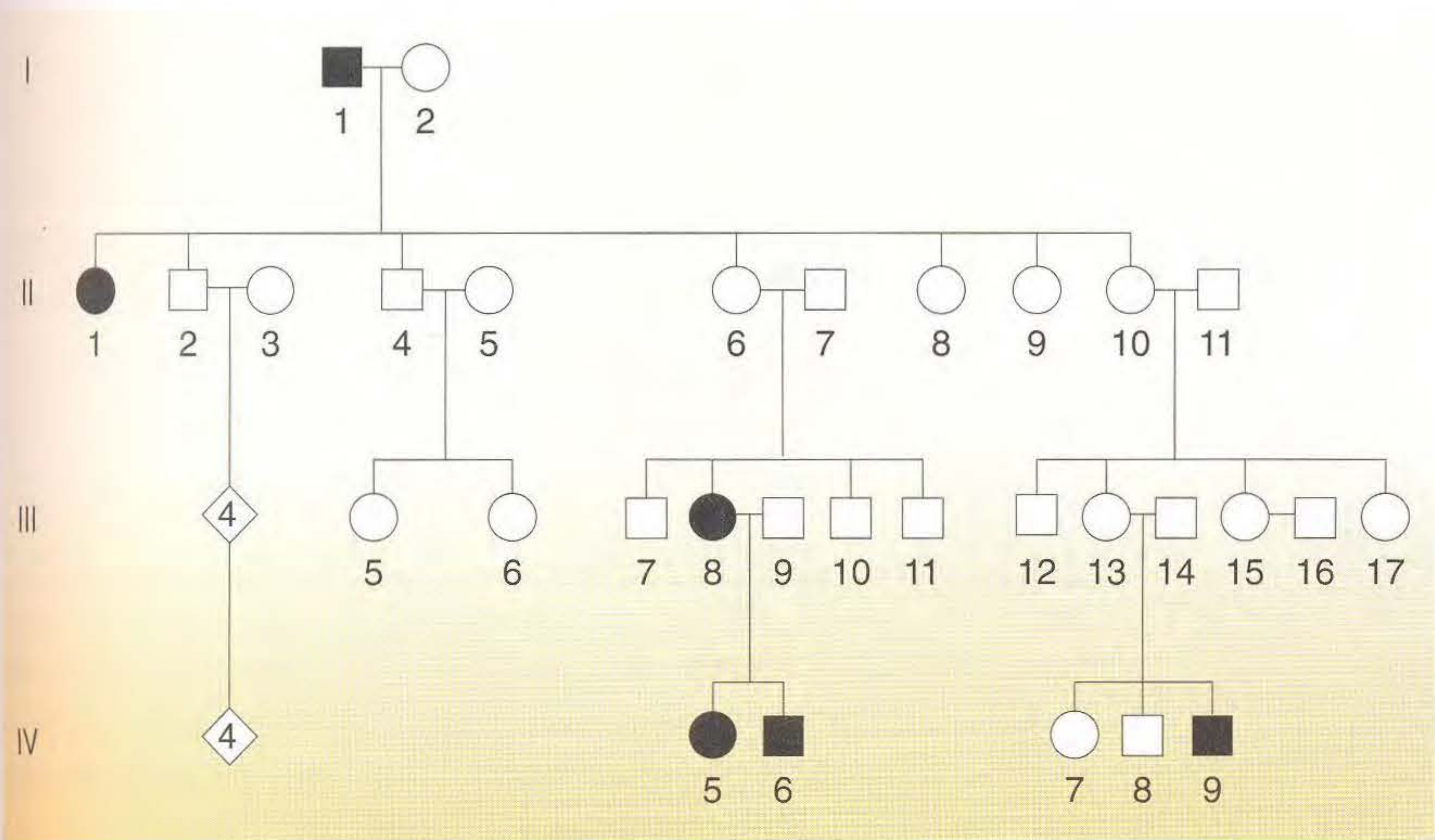
temperatura permissiva

temperatura restritiva

teste de complementação

Problemas Resolvidos

Problema resolvido 1. A maioria dos heredogramas apresenta polidactilia (Figura 2.33) sendo herdada como autossômica dominante rara, mas os heredogramas de algumas famílias não se ajustam totalmente aos padrões esperados para tal herança. Tal heredograma é mostrado aqui. (Os losangos não-sombreados representam os números específicos de pessoas de sexo desconhecido.)



SOLUÇÃO

- a. A expectativa normal para um autossômico dominante é que cada indivíduo afetado tenha um genitor afetado, mas essa expectativa não é vista nesse heredograma, o que constitui a irregularidade. Quais são as explicações possíveis?
- Poderiam alguns casos de polidactilia ser causados por um gene diferente, um que seja um gene dominante ligado ao X? Essa sugestão não é útil, pois ainda teríamos que explicar a ausência da condição nas pessoas II-6 e II-10. Além disso, postulando-se uma herança recessiva, seja autossômica ou ligada ao sexo, isso iria requerer muitas pessoas no heredograma sendo heterozigotas, o que é inapropriado porque a polidactilia é uma condição rara.
- b. Assim, resta-nos a conclusão de que a polidactilia deve, às vezes, ser incompletamente penetrante. Como descrito neste capítulo, alguns indivíduos que têm o genótipo para determinado fenótipo não o expressam. Nesse heredograma, II-6 e II-10 parecem pertencer a essa categoria; eles devem possuir o gene da polidactilia herdado de I-1, pois o transmitem para sua prole.
- c. Como discutido neste capítulo, a supressão ambiental da expressão de um gene pode causar penetrância incompleta, como pode a supressão por outro gene. Para darmos a explica-

- a. Que irregularidade esse heredograma mostra?
- b. Que fenômeno genético esse heredograma ilustra?
- c. Sugira um mecanismo específico de interação gênica que possa produzir tal heredograma, mostrando os genótipos dos membros familiares pertinentes.

ção genética requisitada, devemos ter uma hipótese genética. O que precisamos para explicar? O importante é que I-1 passa o gene para dois tipos de prole, representados por II-1, que expressa o gene, e por II-6 e II-10, que não o expressam. (Por esse heredograma, não podemos dizer se outros filhos de I-1 têm o gene.) A supressão genética está funcionando? I-1 não tem um alelo supressor, pois ele expressa polidactilia. Assim, a única pessoa da qual o supressor pode vir é I-2. Além disso, I-2 deve ser heterozigoto para o gene supressor porque pelo menos um de seus filhos expressa polidactilia. Assim, formulamos a hipótese de que a reprodução na geração I deve ter sido

$$(I-1) P/p \cdot s/s \times (I-2) p/p \cdot S/s$$

onde S é o supressor e P é o alelo responsável pela polidactilia. Por essa hipótese, prevemos que a prole irá compreender os seguintes quatro tipos se os genes se distribuírem:

Genótipo	Fenótipo	Exemplo
$P/p \cdot S/s$	normal (suprimido)	II-6, II-10
$P/p \cdot s/s$	polidactilia	II-1
$p/p \cdot S/s$	normal	
$p/p \cdot s/s$	normal	

Se S é raro, as reproduções de II-6 e II-10 provavelmente dão

Genótipo da prole	Exemplo
$P/p \cdot S/s$	III-13
$P/p \cdot s/s$	III-8
$p/p \cdot S/s$	
$p/p \cdot s/s$	

Não podemos excluir as possibilidades de que II-2 e II-4 tenham o genótipo $P/p \cdot S/s$ e que, por acaso, nenhum de seus descendentes sejam afetados.

Problema resolvido 2. Bezouros de certa espécie podem ter asas verdes, azuis ou turquesa. Bezouros verdes foram selecionados de uma população polimórfica de laboratório e cruzados para determinar a herança da cor da asa. Os cruzamentos e resultados são dados no seguinte quadro:

Cruzamento	Genitores	Prole
1	azul \times verde	toda azul
2	azul \times azul	$\frac{3}{4}$ azul : $\frac{1}{4}$ turquesa
3	verde \times verde	$\frac{3}{4}$ verde : $\frac{1}{4}$ turquesa
4	azul \times turquesa	$\frac{1}{2}$ azul : $\frac{1}{2}$ turquesa
5	azul \times azul	$\frac{3}{4}$ azul : $\frac{1}{4}$ verde
6	azul \times verde	$\frac{1}{2}$ azul : $\frac{1}{2}$ verde
7	azul \times verde	$\frac{1}{2}$ azul : $\frac{1}{4}$ verde $\frac{1}{4}$ turquesa
8	turquesa \times turquesa	toda turquesa

- a. Deduza a base genética da cor da asa nessa espécie.
- b. Escreva os genótipos de todos os genitores e prole o mais completamente possível.

SOLUÇÃO

a. Esses dados a princípio parecem complexos, mas o padrão de herança torna-se claro se considerarmos os cruzamentos um de cada vez. Um princípio geral para resolver tais problemas, como já vimos, é começar observando todos os cruzamentos e agrupando os dados para revelar os padrões.

Um padrão que emerge de uma visão geral dos dados é que todas as proporções são de um gene: não há evidência de dois genes separados tomando parte. Como tal variação pode ser explicada com um só gene? A resposta é que há variação para o próprio gene único, isto é, alelismo múltiplo. Talvez existam três alelos para um gene; chamemos o gene de w (para a cor da asa) e representemos os alelos por w^g, w^b e w^t . Agora temos um problema adicional, que é determinar a dominância desses alelos.

O cruzamento 1 nos diz algo sobre a dominância porque toda a prole de um cruzamento azul \times verde é azul; portanto, azul parece ser dominante em relação a verde. Essa condição é apoiada pelo cruzamento 5, pois o determinante verde deveria estar presente no estoque parental para aparecer na prole. O cruzamento 3 nos informa sobre os determinantes de turquesa, que deveriam estar presentes, embora não expressos, no estoque parental, pois existem asas turquesa na prole. Logo, verde deve ser dominante em relação a turquesa.

Assim, formamos um modelo no qual a dominância é $w^b > w^g > w^t$. De fato, a posição deduzida do alelo w^t na parte inferior da série de dominância é apoiada pelos resultados do cruzamento 7, onde turquesa aparece na prole do cruzamento azul \times verde.

b. Agora é só uma questão de deduzir os genótipos específicos. Note que a pergunta diz que os genitores foram colhidos de uma população polimórfica, o que significa que eles podem ser homozigotos ou heterozigotos. Um genitor com asas azuis, por exemplo, pode ser homozigoto (w^b/w^b) ou heterozigoto (w^b/w^g ou w^b/w^t). Aqui, um pouco de ensaio e erro e bom senso são necessários, mas, nesse estágio, a pergunta foi essencialmente respondida, e o que falta é “cruzar os t e pontuar os i ”. O seguintes genótipos explicam os resultados. Uma barra indica que o genótipo pode ser ou homozigoto ou heterozigoto, e ter um segundo alelo nas séries alélicas.

Cruzamento	Genitores	Prole
1	$w^b/w^b \times w^g/-$	w^b/w^g ou $w^b/-$
2	$w^b/w^t \times w^b/w^t$	$\frac{3}{4} w^b/- : \frac{1}{4} w^t/w^t$
3	$w^g/w^t \times w^g/w^t$	$\frac{3}{4} w^g/- : \frac{1}{4} w^t/w^t$
4	$w^b/w^t \times w^t/w^t$	$\frac{1}{2} w^b/w^t : \frac{1}{2} w^t/w^t$
5	$w^b/w^g \times w^b/w^g$	$\frac{3}{4} w^b/- : \frac{1}{4} w^g/w^g$
6	$w^b/w^g \times w^g/w^g$	$\frac{1}{2} w^b/w^g : \frac{1}{2} w^g/w^g$
7	$w^b/w^t \times w^g/w^t$	$\frac{1}{2} w^b/- : \frac{1}{4} w^g/w^t : \frac{1}{4} w^t/w^t$
8	$w^t/w^t \times w^t/w^t$	toda w^t/w^t

Problema resolvido 3. As folhas de abacaxi podem ser classificadas em três tipos: espinhosa (S), ponta espinhosa (ST) e pequena (não-espinhosa; P). Nos cruzamentos entre as linhagens puras seguidos de entrecruzamentos da F_1 , surgiram os seguintes resultados:

Cruzamento	Genitores	Fenótipos	
		F_1	F_2
1	ST \times S	ST	99 ST : 34 S
2	P \times ST	P	120 P : 39 ST
3	P \times S	P	95 P : 25 ST : 8 S

- a. Atribua símbolos gênicos. Explique esses resultados com relação aos genótipos produzidos e suas proporções.
- b. Usando o modelo da parte a, cite as proporções fenotípicas que você esperaria se fizesse o cruzamento (1) da prole F_1 de pequena \times espinhosa com o estoque parental espinhosa e (2) a prole F_1 de pequena \times espinhosa com a prole F_1 de espinhosa \times ponta espinhosa.

SOLUÇÃO

a. Primeiro, vejamos as proporções de F_2 . Já vimos as proporções 3:1 nos cruzamentos 1 e 2, indicando segregações monogênicas. O cruzamento 3, entretanto, mostra uma proporção que é quase certamente uma proporção 12:3:1. Como sabemos essa proporção? Bem, simplesmente não existem tantas proporções complexas em genética, e o ensaio e erro

nos traz rapidamente a proporção 12:3:1. No total de 128 da prole, são esperados os números 96:24:8, mas os números reais ajustam-se muito bem às expectativas.

Um dos princípios deste capítulo é que as proporções mendelianas modificadas revelam interações gênicas. O cruzamento 3 dá números de F_2 apropriados para uma proporção mendeliana diíbrida modificada, e, assim, parece que estamos lidando com uma interação de dois genes. Parece o lugar mais promissor para começar; podemos voltar aos cruzamentos 1 e 2 e tentar ajustá-los depois.

Qualquer proporção diíbrida é baseada nas proporções fenotípicas 9:3:3:1. Nossos grupos de modificações observadas são:

9 $A/- ; B/-$	} 12 pequena
3 $A/- ; b/b$	
3 $a/a ; B/-$	3 ponta espinhosa
1 $a/a ; b/b$	1 espinhosa

Assim, sem nos preocuparmos com o nome do tipo de interação gênica (isso não está mesmo sendo solicitado), já podemos definir nossos três fenótipos de folha de abacaxi em relação aos pares alélicos propostos A/a e B/b :

pequena	= $A/-$ (B/b irrelevante)
ponta espinhosa	= $a/a ; B/-$
espinhosa	= $a/a ; b/b$

E quanto aos genitores do cruzamento 3? O genitor espinhoso deve ser $a/a ; b/b$, e, como o gene B é necessário para produzir folhas de pontas espinhosas na F_2 , o genitor pequeno deve ser $A/A ; B/B$. (Note que fomos *informados* de que todos os genitores são puros, ou homozigotos.) A F_1 , portanto, deve ser $A/a ; B/b$.

Sem pensarmos muito, podemos escrever o cruzamento 1 como segue:

$$a/a ; B/B \times a/a ; b/b \longrightarrow$$

$$a/a ; B/b \begin{cases} \frac{3}{4} a/a ; B/- \\ \frac{1}{4} a/a ; b/b \end{cases}$$

O cruzamento 2 pode ser parcialmente escrito sem pensarmos muito usando nossos símbolos gênicos arbitrários:

$$A/A ; -/- \times a/a ; B/B \longrightarrow$$

$$A/a ; B/- \begin{cases} \frac{3}{4} A/- ; -/- \\ \frac{1}{4} a/a ; B/- \end{cases}$$

Sabemos que a F_2 do cruzamento 2 mostra segregação monogênica, e parece certo agora que o par de alelos A/a tem um papel. Mas o alelo B é necessário para produzir o fenótipo pontudo; logo, todas as plantas devem ser homozigotas B/B :

$$A/A ; B/B \times a/a ; B/B \longrightarrow$$

$$A/a ; B/B \begin{cases} \frac{3}{4} A/- ; B/B \\ \frac{1}{4} a/a ; B/B \end{cases}$$

Note que as duas segregações monogênicas nos cruzamentos 1 e 2 não mostram que os genes *não* estão interagindo. O que é mostrado é que a interação de dois genes *não* é *revelada* por esses cruzamentos; apenas pelo cruzamento 3, no qual a F_1 é heterozigota para ambos os genes.

b. Agora é uma questão simplesmente de usar as leis de Mendel para prever os resultados do cruzamento:

$$(1) A/a ; B/b \times a/a ; b/b \longrightarrow \left. \begin{array}{l} \frac{1}{4} A/a ; B/b \\ \frac{1}{4} A/a ; b/b \\ \frac{1}{4} a/a ; B/b \\ \frac{1}{4} a/a ; b/b \end{array} \right\} \begin{array}{l} \text{pequena} \\ \text{pequena} \\ \text{ponta espinhosa} \\ \text{espinhosa} \end{array}$$

(distribuição independente em cruzamento-teste padrão)

$$(2) A/a ; B/b \times a/a ; B/b \longrightarrow$$

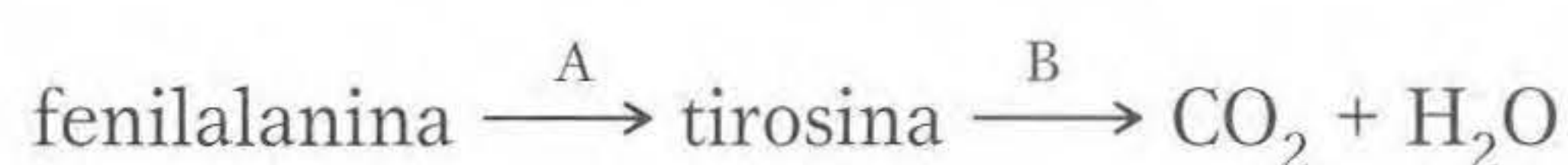
$$\begin{array}{l} \frac{1}{2} A/a \begin{cases} \frac{3}{4} B/- \longrightarrow \frac{3}{8} \\ \frac{1}{4} b/b \longrightarrow \frac{1}{8} \end{cases} \\ \frac{1}{2} a/a \begin{cases} \frac{3}{4} B/- \longrightarrow \frac{3}{8} \\ \frac{1}{4} b/b \longrightarrow \frac{1}{8} \end{cases} \end{array} \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} \frac{1}{2} \text{ pequena} \\ \text{ponta espinhosa} \\ \text{espinhosa} \end{array}$$

Problemas

PROBLEMAS BÁSICOS

- Em humanos, a doença galactosemia causa retardo mental em idade infantil. A lactose (açúcar do leite) é degradada em galactose mais glicose. Normalmente, a galactose é decomposta pela enzima galactose-1-fosfato uridiltransferase (GALT). Entretanto, em pacientes com galactosemia, GALT é inativa, levando ao acúmulo de altos níveis de galactose, a qual, no cérebro, causa retardo mental. Como você faria uma cura secundária para galactosemia? Você esperaria que esse fenótipo de doença fosse dominante ou recessivo?
- Em humanos, a PKU (fenilcetonúria) é uma doença causada por uma ineficiência enzimática na etapa A da sequência simplificada de reações, e a AKU (alcaptonú-

ria) é devida a uma ineficiência enzimática em uma das etapas, resumida aqui como etapa B:



Uma pessoa com PKU casa-se com uma pessoa com AKU. Que fenótipos você espera para seus filhos? Todos normais, todos tendo só PKU, todos tendo só AKU, todos tendo PKU e AKU ou alguns tendo AKU e outros tendo PKU?

- Em *Drosophila*, o autossômico recessivo *bw* causa olho castanho-escuro, e o autossômico recessivo não-ligado *st* causa um olho escarlate-claro. Um homozigoto para ambos os genes tem olho branco. Assim, temos as seguintes correspondências entre genótipos e fenótipos:

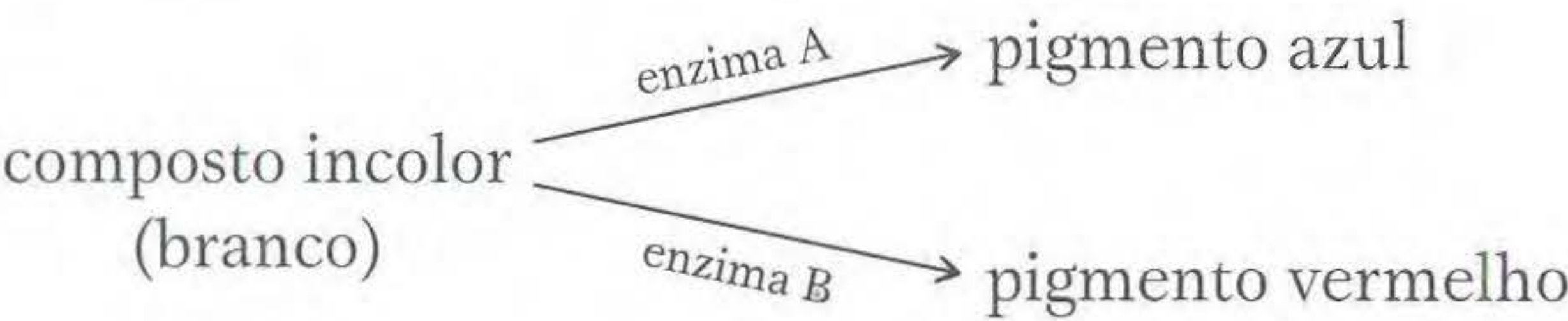
$st^+/st^+; bw^+/bw^+$ = olho vermelho (tipo selvagem)
 $st^+/st^+; bw/bw$ = olho castanho
 $st/st; bw^+/bw^+$ = olho escarlata
 $st/st; bw/bw$ = olho branco

Construa uma via biossintética hipotética mostrando como os produtos gênicos interagem e por que combinações mutantes diferentes têm fenótipos diferentes.

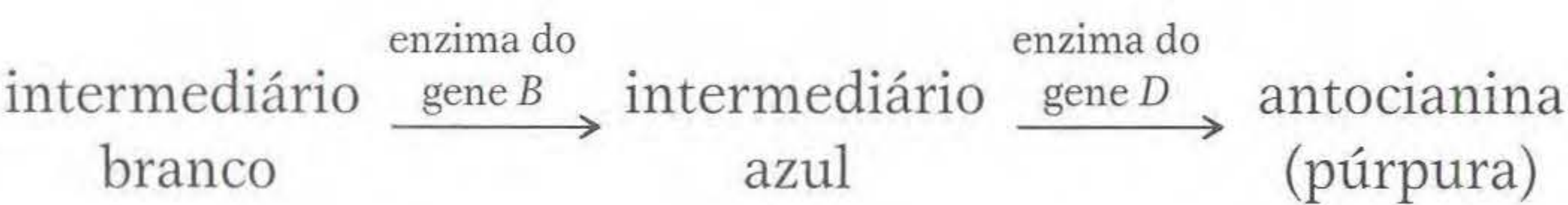
4. Foram isolados vários mutantes, todos os quais requerem o composto G para crescimento. Os compostos (A a E) na via biossintética para G são conhecidos, mas sua ordem na via não é conhecida. Cada composto é testado quanto à sua habilidade em apoiar o crescimento de cada mutante (1 a 5). No quadro seguinte, um sinal + indica crescimento e um sinal - indica não-crescimento.

	Composto testado					
	A	B	C	D	E	G
Mutante 1	-	-	-	+	-	+
2	-	+	-	+	-	+
3	-	-	-	-	-	+
4	-	+	+	+	-	+
5	+	+	+	+	-	+

- a. Qual a ordem dos compostos de A a E na via?
- b. Em que ponto na via cada mutante está bloqueado?
- c. Um heterocácion composto de mutantes duplos 1,3 e 2,4 cresceria em meio mínimo? De 1,3 e 3,4 cresceria? De 1,2 e 2,4 e 1,4 cresceria?
- 5. Em certa planta, as pétalas das flores são normalmente púrpura. Surgem duas mutações recessivas em plantas separadas, e são encontradas estando em cromossomos diferentes. A mutação 1 (m_1) dá pétalas azuis quando em homozigose (m_1/m_1). A mutação 2 (m_2) dá pétalas vermelhas quando em homozigose (m_2/m_2). Os bioquímicos trabalhando na síntese de pigmentos de flor nessa espécie já descreveram a seguinte via:



- a. Que mutante você esperaria como sendo deficiente na atividade da enzima A?
- b. Uma planta tem o genótipo $M_1/m_1; M_2/m_2$. Qual você esperaria que fosse o fenótipo?
- c. Se a planta na parte b for autofecundada, que cores da prole você esperaria e em que proporções?
- d. Por que esses mutantes são recessivos?
- 6. Em ervilhas, a síntese de pigmento púrpura antocianina nas pétalas é controlado por dois genes, B e D. A via é



- a. Que cores de pétalas você esperaria em uma planta pura (homozigota) incapaz de catalisar a primeira reação?

- b. Que cor de pétalas você esperaria em uma planta pura incapaz de catalisar a segunda reação?
- c. Se as plantas nas partes a e b forem cruzadas, que cor de pétalas as plantas de F_1 terão?
- d. Que proporção de plantas púrpura:azul:branca você esperaria na F_2 ?
- 7. Se um homem de grupo sanguíneo AB casa-se com uma mulher do grupo sanguíneo A cujo pai era do grupo O, que grupos sanguíneos diferentes esse homem e essa mulher esperam que seus filhos tenham?
- 8. A maioria da plumagem de galinhas erminette é de cor clara, com uma ocasionalmente preta, dando um aspecto salpicado. Um cruzamento de erminettes produziu um total de 48 na prole, sendo 22 erminettes, 14 pretas e 12 brancas puras. Que base genética é sugerida para o padrão erminette? Como você pode testar sua hipótese?
- 9. Os rabanetes podem ser longos, redondos ou ovais, e podem ser vermelhos, brancos ou púrpura. Você cruza uma variedade longa, branca, com uma redonda, vermelha e obtém uma F_1 oval, púrpura. A F_2 mostra as nove classes fenotípicas seguintes: 9 longas, vermelhas; 15 longas, púrpura; 19 ovais, vermelhas; 32 ovais, púrpura; 8 longas, brancas; 16 redondas, púrpura; 8 redondas, brancas; 16 ovais, brancas; e 9 redondas, vermelhas.
 - a. Forneça uma explicação genética para esses resultados. Esteja certo em definir os genótipos e mostrar a constituição dos genitores, a F_1 e a F_2 .
 - b. Preveja as proporções genotípicas e fenotípicas na prole de um cruzamento entre um rabanete longo, púrpura e um oval, púrpura.
- 10. Nas séries de alelos múltiplos que determinam a cor da pelagem em coelhos, c^+ codifica aguti, c^{ch} codifica chinchila (uma pelagem bege) e c^h codifica himalaio. A dominância é na ordem $c^+ > c^{ch} > c^h$. Em um cruzamento de $c^+/c^{ch} \times c^{ch}/c^h$, que proporção da prole será chinchila?
- 11. Preto, sépia, creme e albino são cores de pelagem de cobaias (porquinho-da-índia). Animais individuais (não necessariamente de raças puras) apresentando essas cores foram entrecruzados; os resultados são tabulados como em seguida, onde as abreviações A (albino), B (preto), C (creme) e S (sépia) representam os fenótipos:

Cruzamento	Fenótipos parentais	Fenótipos da prole			
		B	S	C	A
1	B × B	22	0	0	7
2	B × A	10	9	0	0
3	C × C	0	0	34	11
4	S × C	0	24	11	12
5	B × A	13	0	12	0
6	B × C	19	20	0	0
7	B × S	18	20	0	0
8	B × S	14	8	6	0
9	S × S	0	26	9	0
10	C × A	0	0	15	17

a. Deduza a herança dessas cores de pelagem e use símbolos gênicos de sua própria escolha. Mostre todos os genótipos parentais e da prole.

b. Se os animais pretos nos cruzamentos 7 e 8 forem cruzados, que proporções de prole você prevê usando seu modelo?

12. Em uma maternidade, quatro crianças casualmente foram misturadas. Os tipos ABO das quatro crianças são conhecidos como sendo O, A, B e AB. Os tipos ABO dos quatro grupos de genitores foram determinados. Indique que criança pertence a cada um dos conjuntos de genitores: (a) $AB \times O$, (b) $A \times O$, (c) $A \times AB$, (d) $O \times O$.

13. Considere dois polimorfismos sanguíneos que os humanos têm além do sistema ABO. Dois alelos L^M e L^N determinam os grupos sanguíneos M, N, e MN. O alelo dominante R de um gene diferente faz com que a pessoa tenha o fenótipo Rh^+ (rhesus positivo), enquanto o homozigoto para r é Rh^- (rhesus negativo). Dois homens levam aos tribunais uma disputa de paternidade, cada um dizendo que três crianças lhe pertencem. Os grupos sanguíneos dos homens, crianças e suas mães eram os seguintes:

Pessoa	Grupo sanguíneo		
marido	O	M	Rh^+
amante da esposa	AB	MN	Rh^-
esposa	A	N	Rh^+
criança 1	O	MN	Rh^+
criança 2	A	N	Rh^+
criança 3	A	MN	Rh^-

Por essa evidência, pode ser estabelecida a paternidade das crianças?

14. Em um rancho de raposas em Wisconsin, surgiu uma mutação que deu uma cor de pelagem platina. A cor platina demonstrou-se muito popular para os compradores de casacos de pele de raposa, mas os criadores não conseguiram desenvolver uma linhagem pura de platina. Toda vez que duas platinas eram cruzadas, apareciam na prole algumas raposas normais. Por exemplo, as reproduções repetidas do mesmo par de platina produzia 82 platinas e 38 normais na prole. Todos os outros desses cruzamentos deram proporções de pelagem similares. Formule uma hipótese genética concisa que responda por esses resultados.

15. Por um período de sete anos, Hans Nachtsheim investigou uma anomalia herdada dos leucócitos de coelhos. Essa anomalia, chamada de *anomalia Pelger*, é a parada de segmentação de alguns leucócitos. Essa anomalia não parece ser um grave inconveniente para os coelhos.

a. Quando coelhos apresentando a típica anomalia de Pelger foram cruzados com coelhos de um estoque normal, Nachtsheim contou 217 proles apresentando a anomalia Pelger e 237 normais na prole. Qual parece ser a base genética da anomalia Pelger?

b. Quando coelhos com a anomalia Pelger foram cruzados uns com os outros, Nachtsheim encontrou 223 proles

normais, 439 apresentando a anomalia Pelger, e 39 extremamente anormais. Essa prole muito anormal não só tinha leucócitos defeituosos, mas também apresentava graves deformidades do sistema esquelético; quase todos eles morreram logo após o nascimento. Em termos genéticos, o que você supõe que esses coelhos extremamente defeituosos representam? Por que você supõe que havia apenas 39 deles?

c. Que evidência experimental adicional você colheria para apoiar ou desaprovar suas respostas para a parte b?

d. Em Berlim, cerca de 1 humano em cada 1.000 apresenta uma anomalia de Pelger dos leucócitos muito similar a essa descrita para os coelhos. A anomalia é herdada como dominante simples, mas o tipo homozigoto não foi observado em humanos. Você pode sugerir o motivo, fazendo uma analogia com a condição em coelhos?

e. Novamente por analogia com os coelhos, que fenótipos e genótipos devem ser esperados entre os filhos de um casal, ambos com a anomalia de Pelger?

(O problema 15 é de A.M. Srb, R.D. Owen, and R.S. Edgar, *General Genetics*, 2nd ed. W.H. Freeman and Company, 1965.)

16. Foram cruzadas duas moscas-das-frutas de aspecto normal, e, na prole, havia 202 fêmeas e 98 machos.

a. O que há de incomum nesse resultado?

b. Forneça uma explicação genética para essa anomalia.

c. Forneça um teste para sua hipótese.

17. Você recebeu uma fêmea virgem de *Drosophila*. Você nota que as cerdas de seu tórax são muito mais curtas que o normal. Você a cruza com um macho normal (com cerdas longas) e obtém a seguinte prole F_1 : $\frac{1}{3}$ de fêmeas com cerdas curtas, $\frac{1}{3}$ de fêmeas com cerdas longas e $\frac{1}{3}$ de machos com cerdas longas. Um cruzamento de fêmeas da F_1 de cerdas longas com seus irmãos só deu uma F_2 de cerdas longas. Um cruzamento de fêmeas de cerdas curtas com seus irmãos deu $\frac{1}{3}$ de fêmeas de cerdas curtas, $\frac{1}{3}$ de fêmeas de cerdas longas e $\frac{1}{3}$ de machos de cerdas longas. Forneça uma hipótese genética que explique esses resultados, mostrando os genótipos de cada cruzamento.

18. Um alelo dominante H reduz o número de cerdas do corpo que as *Drosophilas* têm, dando origem ao fenótipo "sem cerdas". Na condição homozigota, H é letal. Um alelo dominante S que se segrega independentemente não tem efeito no número de cerdas, exceto na presença de H , em cujo caso uma única dose de S suprime o fenótipo sem cerdas, restaurando assim o fenótipo com cerdas. Entretanto, S também é letal na condição homozigota (S/S).

a. Que proporção de moscas com cerdas para sem cerdas você encontraria na prole viva de um cruzamento entre duas moscas com cerdas, ambas levando H na condição suprimida?

b. Quando a prole sem cerdas é retrocruzada com uma mosca parental com cerdas, que proporção fenotípica você esperaria encontrar entre a prole viva?

19. Após irradiar células tipo selvagem de *Neurospora* (um fungo haplóide), um geneticista encontra dois mutan-

tes auxotróficos que precisam de leucina. Ele combina os dois mutantes em um heterocáron e descobre que o heterocáron é prototrófico.

a. As mutações nos dois auxotróficos estão no *mesmo* gene na via para a síntese de leucina ou em dois genes *diferentes* nessa via? Explique.

b. Escreva o genótipo das duas linhagens de acordo com seu modelo.

c. Que prole e em que proporções você prevê crossing dos dois mutantes auxotróficos? (Suponha distribuição independente.)

20. Um geneticista de leveduras irradia células haplóides de uma linhagem que é um mutante auxotrófico precisando de adenina, causada por mutação no gene *ade1*. Milhões de células irradiadas são plaqueadas em meio mínimo, e um pequeno número de células divide-se e produz colônias auxotróficas. Essas colônias são cruzadas individualmente com uma linhagem tipo selvagem. São obtidos dois tipos de resultados:

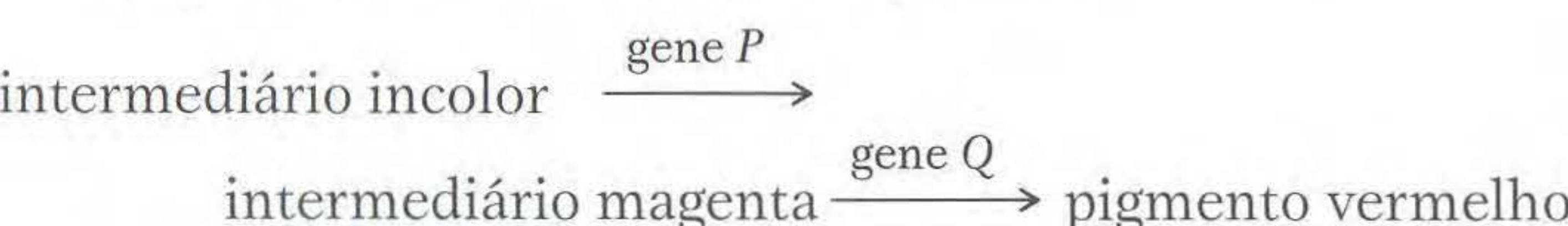
- (1) prototrófico × tipo selvagem: prole toda prototrófica
- (2) prototrófica × tipo selvagem: prole 75% prototrófica, 25% auxotrófica precisando de adenina.

a. Explique a diferença entre esses dois tipos de resultados.

b. Escreva os genótipos dos prototróficos em cada caso.

c. Que fenótipos da prole e proporções você prevê para o cruzamento de um prototrófico do tipo 2 com o auxotrófico original *ade1*?

21. Em rosas, a síntese do pigmento vermelho é por duas etapas em uma via, como a seguinte:



a. Qual seria o fenótipo de uma planta homozigota para uma mutação nula do gene *P*?

b. Qual seria o fenótipo de uma planta homozigota para uma mutação nula do gene *Q*?

c. Qual seria o fenótipo de uma planta homozigota para mutações nulas dos genes *P* e *Q*?

d. Escreva os genótipos de três linhagens nas partes a, b e c.

e. Que proporção de *F*₂ é esperada do cruzamento de plantas das partes a e b? (Suponha distribuição independente.)

22. Como a boca-de-leão (*Antirrhinum*) possui o pigmento antocianina, ela tem pétalas de cor púrpura avermelhada. Duas linhagens puras sem antocianina de *Antirrhinum* foram desenvolvidas, uma na Califórnia e outra na Holanda. Elas parecem idênticas por não terem nenhum pigmento vermelho, o que se manifesta com flores brancas (albinas). Entretanto, quando as pétalas de duas linhagens foram cultivadas juntas em tampão no mesmo tubo de ensaio, a solução, que a princípio parecia incolor, gradualmente ficou vermelha.

a. Que experimentos-controle um pesquisador deve fazer antes de continuar com outras análises?

b. O que poderia explicar a produção da cor vermelha no tubo de ensaio?

c. De acordo com sua explicação para a parte b, quais seriam os genótipos das duas linhagens?

d. Se as duas linhagens brancas forem cruzadas, qual seria sua previsão para os fenótipos de *F*₁ e *F*₂?

23. A galinha frizzle é muito admirada pelos criadores. Ela tem seu nome devido ao modo incomum como suas penas se curvam para cima, dando a impressão de que (nas memoráveis palavras do geneticista animal F.B. Hutt) “foram dobradas para trás por um tipo nó de madeira”. Infelizmente, frizzle não é pura; quando duas frizzles são entrecruzadas, elas sempre produzem 50% frizzles, 25% normais e 25% com penas peculiares lanosas que logo caem, deixando as aves nuas.

a. Cite uma explicação genética para esses resultados, mostrando os genótipos e todos os fenótipos, e diga como sua explicação funciona.

b. Se você quer uma produção em massa de frizzle para vender, que tipos seriam melhores como par de cruzamento?

24. As pétalas da planta *Collinsia parviflora* são normalmente azuis, dando a espécie seu nome comum *blue-eyed Mary*. As duas linhagens puras de cruzamento foram obtidas de variantes de cor encontradas na natureza; a primeira linha tinha pétalas rosa, e a segunda linha tinha pétalas brancas. Foram feitos os seguintes cruzamentos entre linhagens puras, com os resultados mostrados:

Genitores	F ₁	F ₂
azul × branca	azul	101 azuis, 33 brancas
azul × rosa	azul	192 azuis, 63 rosa
rosa × branca	azul	272 azuis, 121 brancas, 89 rosa

a. Explique geneticamente esses resultados. Defina os símbolos alélicos que você usa e mostre a constituição genética dos genitores, da *F*₁ e da *F*₂ em cada cruzamento.

b. Um cruzamento entre certa planta *F*₂ azul e outra *F*₂ branca deu uma prole na qual ³/₈ eram azuis, ¹/₈ rosa e ¹/₈ de cor branca. Que genótipos essas duas plantas *F*₂ tiveram?



Solucionando o Prolema 24

- Qual a característica que está sendo estudada?
- Qual é o fenótipo tipo selvagem?
- O que é uma variante?
- Quais são as variantes neste problema?
- O que significa “na natureza”?
- De que modo as variantes foram encontradas na natureza? (Descreva a cena.)
- Em que estágios no experimento as sementes seriam usadas?
- O modo de escrever um cruzamento “azul × branco”, por exemplo, significa o mesmo que “branco × azul”? Você espera resultados similares? Por que sim ou por que não?

9. De que modo as duas fileiras no quadro diferem da terceira?
10. Que fenótipos são dominantes?
11. O que é complementação?
12. De onde vem o azul na prole do cruzamento rosa × branco?
13. Que fenômeno genético a produção de uma F₁ azul de genitores rosa e branco representa?
14. Cite quaisquer proporções que você possa ver.
15. Existem algumas proporções monoíbridas?
16. Existem proporções diíbridas?
17. O que a observação de proporções monoíbridas e diíbridas lhe diz?
18. Cite quatro proporções mendelianas modificadas que você pode pensar.
19. Existem proporções mendelianas modificadas no problema?
20. O que proporções mendelianas modificadas em geral indicam?
21. O que é indicado por proporção modificada específica ou proporções neste problema?
22. Desenhe cromossomos que representam meioses nos genitores do cruzamento azul × branco e representam meioses na F₁.
23. Repita a etapa 22 no cruzamento azul × rosa.
25. Uma mulher que tem um poodle albino puro (um fenótipo autossômico recessivo) deseja filhotes brancos; assim, ela levou o cão a um criador, que disse que iria cruzar a fêmea com um macho albino, também de um estoque puro. Quando nasceram os seis filhotes, todos eram pretos; e, assim, a mulher processou o criador, dizendo que ele substituiu o macho por um cão preto, dando os seis filhotes indesejados. Você foi chamado como especialista testemunha, e a defesa lhe pergunta se é possível produzir prole preta de dois albinos recessivos puros. Que testemunho você daria?
26. Uma planta boca-de-leão pura para pétalas brancas foi cruzada com uma planta pura para pétalas púrpura, e toda a F₁ tinha pétalas brancas. A F₁ foi autofecundada. Entre a F₂, foram observados três fenótipos nos seguintes números:

branca	240
púrpura uniforme	61
púrpura pontilhada	19
Total	320

a. Proponha uma explicação para esses resultados, mostrando os genótipos de todas as gerações (faça e explique seus símbolos).

b. Uma planta branca de F₂ foi cruzada com uma púrpura uniforme de F₂, e a prole foi

branca	50%
púrpura uniforme	25%
púrpura pontilhada	25%

Quais eram os genótipos das plantas F₂ cruzadas?

27. A maioria dos bezouros da farinha são pretos, mas muitas variantes de cor são conhecidas. Os cruzamentos de genitores puros produziram os seguintes resultados (veja quadro) na geração F₁, e, inter cruzando a F₁ de cada cruzamento, temos as proporções mostradas para a geração F₂. Os fenótipos estão abreviados Bl, preto; Br, marrom; Y, amarelo; e W, branco.

Cruzamento	Genitores	F ₁	F ₂
1	Br × Y	Br	3 Br : 1 Y
2	Bl × Br	Bl	3 Bl : 1 Br
3	Bl × Y	Bl	3 Bl : 1 Y
4	W × Y	Bl	9 Bl : 3 Y : 4 W
5	W × Br	Bl	9 Bl : 3 Br : 4 W
6	Bl × W	Bl	9 Bl : 3 Y : 4 W

a. A partir desses resultados, deduza e explique a herança dessas cores.

b. Escreva os genótipos de cada um dos genitores, a F₁ e a F₂ em todos os cruzamentos.

28. Dois albinos casam-se e têm quatro filhos normais. Como isso é possível?

29. Considere a produção da cor da flor em ipoméia japonesa (*Pharbitis nil*). Os alelos dominantes de dois genes separados (A/- · b/b ou a/a · B/-) produzem pétalas púrpura. A/- · B/- produz pétalas azuis e a/a · b/b produz pétalas escarlate. Deduza os genótipos dos genitores e a prole nos seguintes cruzamentos:

Cruzamento	Genitores	Prole
1	azul × escarlate	$\frac{1}{4}$ azul : $\frac{1}{2}$ púrpura : $\frac{1}{4}$ escarlate
2	púrpura × púrpura	$\frac{1}{4}$ azul : $\frac{1}{2}$ púrpura : $\frac{1}{4}$ escarlate
3	azul × azul	$\frac{3}{4}$ azuis : $\frac{1}{4}$ púrpura
4	azul × púrpura	$\frac{3}{8}$ azuis : $\frac{4}{8}$ púrpura : $\frac{1}{8}$ escarlate
5	púrpura × escarlate	$\frac{1}{2}$ púrpura : $\frac{1}{2}$ escarlate

30. Os agricultores de milho obtiveram linhagens puras cujos grãos ficavam vermelho-sol, rosa, escarlate ou laranja quando expostos à luz solar (normalmente, os grãos ficam amarelos à luz solar). Alguns cruzamentos produzem os seguintes resultados. Os fenótipos são abreviados O, laranja; P, rosa; Sc, escarlate; e SR, vermelho-sol.

Cruzamento	Genitores	Fenótipos	
		F ₁	F ₂
1	SR × P	todos SR	66 SR : 20 P
2	O × SR	todos SR	998 SR : 314 O
3	O × P	todos O	1.300 O : 429 P
4	O × Sc	todos Y	182Y : 80 O : 58 Sc

Analise os resultados de cada cruzamento e forneça uma hipótese unificadora que explique *todos* os resultados. (Explique todos os símbolos que você usa.)

31. Muitos tipos de animais selvagens têm o padrão de cor aguti, em que cada pêlo tem uma banda amarela ao redor de si.
- a. Camundongos pretos e outros animais pretos não têm a banda amarela; cada um de seus pêlos é todo preto. Essa ausência de aguti selvagem é chamada de *não-aguti*. Quando camundongos da linhagem aguti pura são cruzados com não-aguti, a F_1 é toda aguti e a F_2 tem uma proporção de 3:1 de agutis para não-agutis. Faça um diagrama desse cruzamento, sendo A o representante do alelo responsável pelo fenótipo aguti e a , não-aguti. Mostre os fenótipos e genótipos dos genitores, seus gametas, a F_1 , seus gametas e a F_2 .
- b. Outro desvio de cor herdado em camundongos substitui marrom pela cor preta na pelagem tipo selvagem. Tais camundongos marrom-aguti são chamados *cinnamons*. Quando camundongos tipo selvagem são cruzados com cinnamons, toda a F_1 é tipo selvagem, e a F_2 tem uma proporção 3:1 de tipo selvagem para cinnamon. Faça um diagrama desse cruzamento como na parte a, sendo B o alelo preto tipo selvagem e b o alelo cinnamon marrom.
- c. Quando camundongos de uma linhagem pura cinnamon são cruzados com camundongos de uma linhagem pura não-aguti (preta), toda a F_1 é tipo selvagem. Use um diagrama genético para explicar esse resultado.
- d. Na F_2 do cruzamento da parte c, apareceu uma quarta cor chamada *chocolate*, além da cinnamon parental e não-aguti e do tipo selvagem da F_1 . Os camundongos chocolate têm uma cor marrom uniforme. Qual a constituição genética dos camundongos chocolate?
- e. Supondo que os pares alélicos A/a e B/b distribuíam-se independentemente um do outro, como você espera que sejam as frequências relativas dos quatro tipos de cor na F_2 descrita na parte d? Diagrame o cruzamento das partes c e d, mostrando os genótipos e fenótipos (incluindo os gametas).
- f. Que fenótipos seriam observados e em que proporções na prole de um retrocruzamento de camundongos F_1 da parte c com o estoque parental cinnamon? Com o estoque parental não-aguti (preto)? Diagrame esses retrocruzamentos.
- g. Diagrame um cruzamento-teste para a F_1 da parte c. Que cores resultariam e em que proporções?
- h. Os camundongos albinos (olhos cor-de-rosa) são homozigotos para um membro recessivo de um par alélico C/c , que se segrega independentemente dos pares A/a e B/b . Suponha que você tenha quatro linhagens albinas altamente endocruzadas (e, portanto, supostamente homozigotas). Você cruza cada uma dessas linhagens com uma linhagem tipo selvagem pura, e você cria uma grande prole F_2 de cada cruzamento. Que genótipos para as linhagens albinas você pode deduzir dos seguintes fenótipos F_2 ?

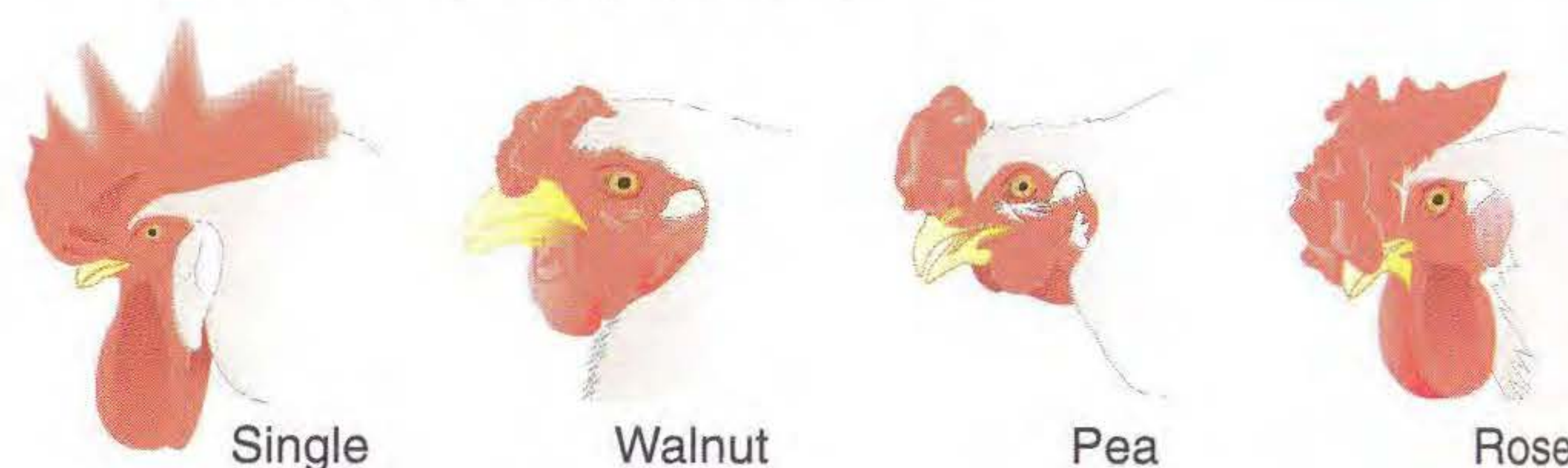
Fenótipos da prole

F_2 da linhagem	Tipo selvagem	Preto	Cinnamon	Chocolate	Albino
1	87	0	32	0	39
2	62	0	0	0	18
3	96	30	0	0	41
4	287	86	92	29	164

(O problema 31 é adaptado de A.M. Srb, R.D. Owen, and R.S. Edgar, *General Genetics*, 2nd ed. W. H. Freeman and Company, 1965.)



32. Um alelo A que não é letal quando homozigoto faz com que os ratos tenham pelagem amarela. O alelo R de um gene separado que segrega independentemente produz uma pelagem preta. Juntos, A e R produzem uma pelagem cinza, enquanto a e r produzem uma pelagem branca. Um macho cinza é cruzado com uma fêmea amarela, e a F_1 é $\frac{3}{8}$ amarela, $\frac{3}{8}$ cinza, $\frac{1}{8}$ preta e $\frac{1}{8}$ branca. Determine os genótipos dos genitores.
33. O genótipo r/r ; p/p dá uma crista *single* a galinhas, $R/-$; $P/-$ dá uma crista *walnut*, r/r ; $P/-$ dá uma crista *pea* e $R/-$; p/p dá uma crista *rose* (veja as ilustrações). Suponha distribuição independente.



- a. Que tipos de cristas aparecerão na F_1 e na F_2 e em que proporções se as aves com cristas *single* são cruzadas com aves de uma linhagem *walnut* pura?
- b. Quais são os genótipos dos genitores em um cruzamento *walnut* \times *rose* do qual a prole é $\frac{3}{8}$ *rose*, $\frac{3}{8}$ *walnut*, $\frac{1}{8}$ *pea* e $\frac{1}{8}$ *single*?
- c. Quais são os genótipos dos genitores em um cruzamento *walnut* \times *rose* do qual toda a prole é *walnut*?
- d. Como muitos genótipos produzem um fenótipo *walnut*? Escreva-os.
34. A produção da cor de olho em *Drosophila* requer um alelo dominante A . O alelo dominante P do segundo gene independente transforma o pigmento em púrpura, mas seu alelo recessivo é vermelho. Uma mosca que não produz pigmento tem olhos brancos. Foram cruzadas duas linhagens puras com os seguintes resultados:
- P fêmea de olho vermelho \times macho de olho branco
- ↓
- F_1 fêmeas de olho púrpura
machos de olho vermelho
- $F_1 \times F_1$
- ↓
- F_2 tanto machos quanto fêmeas: $\frac{3}{8}$ olhos púrpura
 $\frac{3}{8}$ olhos vermelhos
 $\frac{2}{8}$ olhos brancos

Explique esse modo de herança e mostre os genótipos dos genitores, da F₁ e da F₂.

35. Quando cães marrons puros são cruzados com alguns cães brancos puros, todos os filhotes da F₁ são brancos. A prole F₂ de alguns cruzamentos F₁ × F₁ era de 118 filhotes brancos, 32 pretos e 10 marrons. Qual a base genética para esses resultados?

36. Linhagens tipo selvagem do fungo haplóide *Neurospora* podem fazer seu próprio triptofano. Um alelo anormal *td* torna o fungo incapaz de fazer seu próprio triptofano. Um indivíduo com o genótipo *td* só cresce quando seu meio fornece triptofano. O alelo *su* distribui-se independentemente de *td*; seu único efeito é suprimir o fenótipo *td*. Portanto, as linhagens que levam tanto *td* quanto *su* não requerem triptofano para crescimento.

a. Se uma linhagem *td* ; *su* é cruzada com uma linhagem genotipicamente tipo selvagem, que genótipos são esperados na prole e em que proporções?

b. Qual será a proporção de dependentes de triptofano para independentes de triptofano na prole do cruzamento na parte a?

37. Camundongos com os genótipos *A/A* ; *B/B* ; *C/C* ; *D/D* ; *S/S* e *a/a* ; *b/b* ; *c/c* ; *d/d* ; *s/s* são cruzados. A prole é entrecruzada. Que fenótipos serão produzidos na F₂ e em que proporções? (Os símbolos alélicos são *A* = aguti, *a* = uniforme (não-aguti); *B* = pigmento preto, *b* = marrom; *C* = pigmentado, *c* = albino; *D* = não-diluição, *d* = diluição (cor de leite); *S* = não-pontilhado, *s* = pontilhados pigmentados sobre fundo branco.)

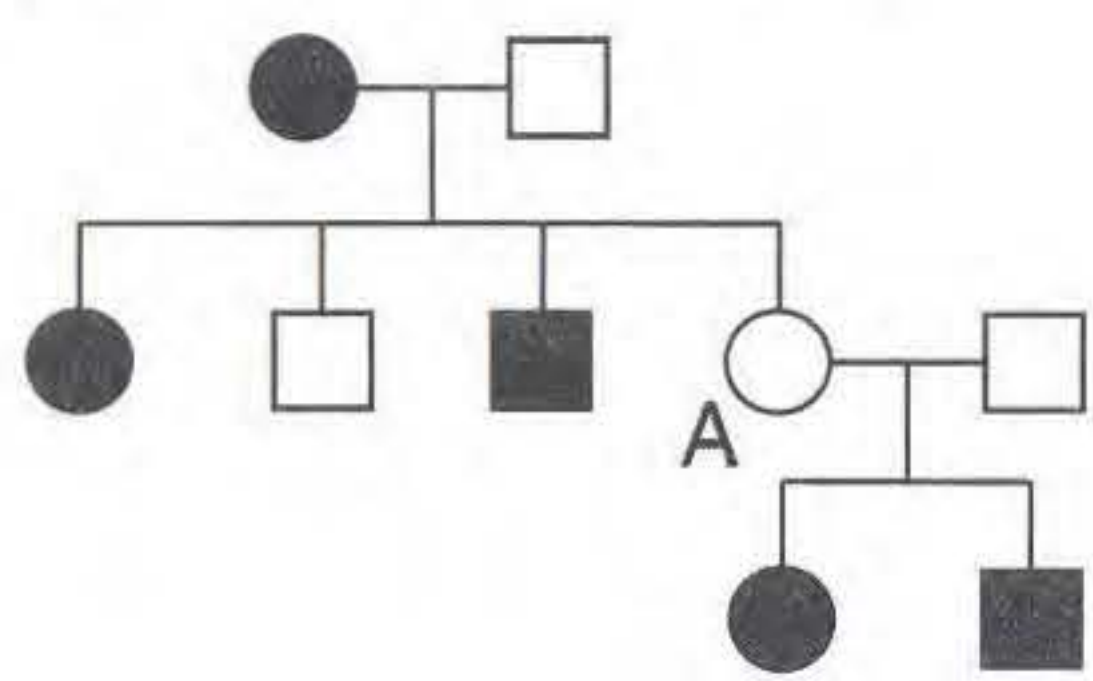
38. Considere os genótipos de duas linhagens de galinhas: a linhagem pura salpicada Honduras é *i/i* ; *D/D* ; *M/M* ; *W/W*, e as da linhagem pura *leghorn* é *I/I* ; *d/d* ; *m/m* ; *w/w*, onde

- I* = plumagem branca, *i* = plumagem colorida
- D* = crista dupla, *d* = crista simples
- M* = barba, *m* = imberbe
- W* = pele branca, *w* = pele amarela

Esses quatro genes distribuem-se independentemente. Começando com essas linhagens puras, qual o modo mais rápido e mais conveniente de gerar uma linhagem pura que tenha plumagem colorida, crista simples, seja imberbe e com pele amarela? Certifique-se de que você mostra

- a. o pedigree do cruzamento.
- b. o genótipo de cada animal representado.
- c. quantos ovos eclodem de cada cruzamento, e por que esse número.
- d. por que seu esquema é o mais rápido e o mais conveniente.

39. O heredograma seguinte é de um fenótipo dominante governado por um alelo autossômico. O que esse heredograma sugere sobre o fenótipo e o que você pode deduzir sobre o genótipo do indivíduo A?



40. A coloração das pétalas em dedaleira é determinada por três genes. *M* codifica uma enzima que sintetiza antocianina, o pigmento púrpura visto nessas pétalas; *m/m* não produz pigmento, resultando no fenótipo albino com manchas amareladas. *D* é um acentuador de antocianina, resultando em um pigmento mais escuro; *d/d* não acentua. No terceiro locus, *w/w* permite a deposição de pigmento nas pétalas, mas *W* impede a deposição do pigmento exceto nos pontos, e, assim, resulta no fenótipo branco salpicado. Considere os seguintes dois cruzamentos:

Cruzamento	Genitores	Prole
1	Púrpura-escuro × branco com manchas amareladas	$\frac{1}{2}$ púrpura-escuro: $\frac{1}{2}$ púrpura-claro
2	branca com manchas amareladas × púrpura-claro	$\frac{1}{2}$ branca com manchas púrpura: $\frac{1}{4}$ púrpura-escuro: $\frac{1}{4}$ púrpura-claro

Em cada caso, cite os genótipos dos genitores e prole com relação aos três genes.

41. Em uma espécie de *Drosophila*, as asas são normalmente arredondadas, mas você obteve duas linhagens puras, uma com asas ovais e outra com asas afoiçadas. Os cruzamentos entre as linhagens puras revelaram os seguintes resultados:

Genitores		F ₁	
Fêmea	Macho	Fêmea	Macho
afoiçadas	redondas	afoiçadas	afoiçadas
redondas	afoiçadas	afoiçadas	redondas
afoiçadas	ovais	ovais	afoiçadas

a. Cite uma explicação genética para esses resultados, definindo todos os símbolos alélicos.

b. Se as fêmeas F₁ ovais do cruzamento 3 são cruzadas com machos redondos da F₁ do cruzamento 2, que proporções fenotípicas são esperadas para cada sexo na prole?

42. Os camundongos normalmente têm uma banda amarela em cada pêlo, mas são conhecidas variantes com duas ou três bandas. Uma fêmea de camundongo tendo uma banda foi cruzada com um macho tendo três bandas. (Nenhum dos animais eram de uma linhagem pura.) A prole foi

Fêmeas	$\frac{1}{2}$ uma banda	Machos	$\frac{1}{2}$ uma banda
	$\frac{1}{2}$ três bandas		$\frac{1}{2}$ duas bandas

a. Forneça uma explicação clara da herança desses fenótipos.

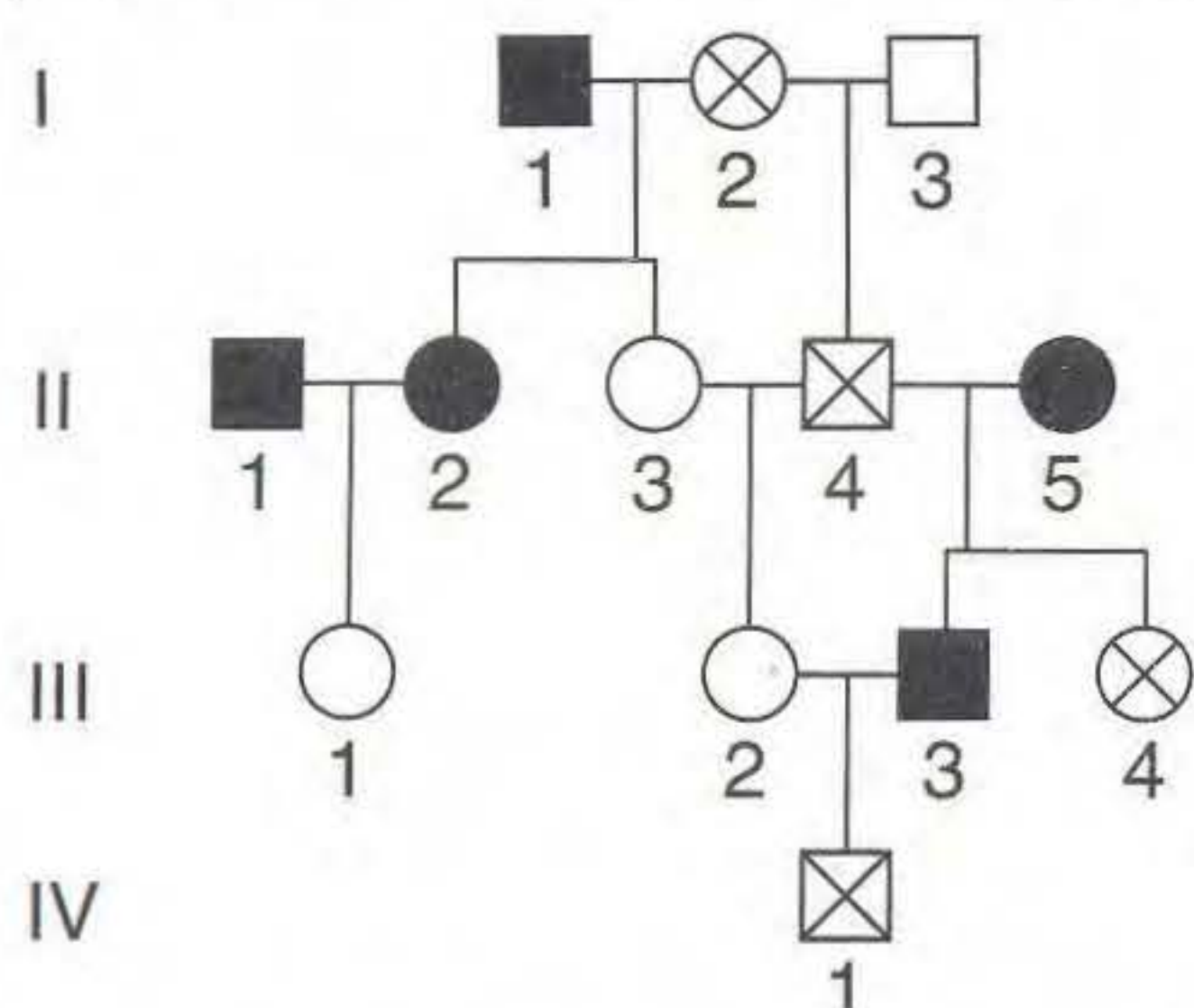
b. De acordo com seu modelo, qual seria o resultado de um cruzamento entre uma filha com três bandas e um filho com uma banda?

43. Em martas, os tipos selvagens têm uma pelagem quase preta. Os criadores desenvolveram muitas linhagens puras de cores variadas para a indústria de casacos de marta. Duas linhagens puras são platina (cinza-azulado)

e aleutian (cinza-metálico). Essas linhagens foram usadas em cruzamentos com os seguintes resultados:

Cruzamento	Genitores	F ₁	F ₂
1	selvagem × platina	selvagem	18 selvagem, 5 platina
2	selvagem × aleutian	selvagem	27 selvagem, 10 aleutian
2	platina × aleutian	selvagem	133 selvagem 41 platina 46 aleutian 17 safira (nova)

- a. Imagine uma explicação genética para esses três cruzamentos. Mostre os genótipos dos genitores, da F₁ e da F₂ nos três cruzamentos, e certifique-se de que você mostrou os alelos de cada gene que você supôs para cada indivíduo.
- b. Preveja as proporções fenotípicas de F₁ e F₂ do cruzamento de safira com platina e linhagens puras aleutians.
44. Em *Drosophila*, um gene autossômico determina a forma das cerdas, com *B* dando retas e *b* inclinadas. Em outro autossomo, há um gene no qual um alelo dominante *I* inibe a formação de cerdas, de modo que a mosca não tem cerdas (*i* não tem efeito fenotípico conhecido).
- a. Se uma mosca de cerdas retas de uma linhagem pura é cruzada com uma mosca de uma linhagem pura sem cerdas, conhecida como tendo um genótipo inibido para inclinada, quais serão os genótipos e fenótipos da F₁ e F₂?
- b. Que cruzamento daria a proporção 4 sem cerdas: 3 retas: 1 inclinada?
45. O pedigree seguinte é relativo a fenótipos oculares em besouros *Tribolium*. Os símbolos escuros representam olhos pretos, os símbolos vazados representam olhos castanhos e os símbolos em X representam o fenótipo “sem olhos”, no qual os olhos estão totalmente ausentes.



- a. A partir desses dados, deduza o modo de herança dos três fenótipos.
- b. Usando símbolos gênicos definidos, mostre o genótipo do besouro II-3.
46. Uma planta tida como heterozigota para um par de alelos *B/b* (onde *B* codifica amarelo e *b* codifica bronze) foi autofecundada, e, na prole, havia 280 plantas amarelas e 120 bronze. Esses resultados apóiam a hipótese de que a planta é *B/b*?
47. Uma planta tida como sendo heterozigota para dois genes que se segregam independentemente (*P/p* ; *Q/q*) foi autofecundada, e a prole era

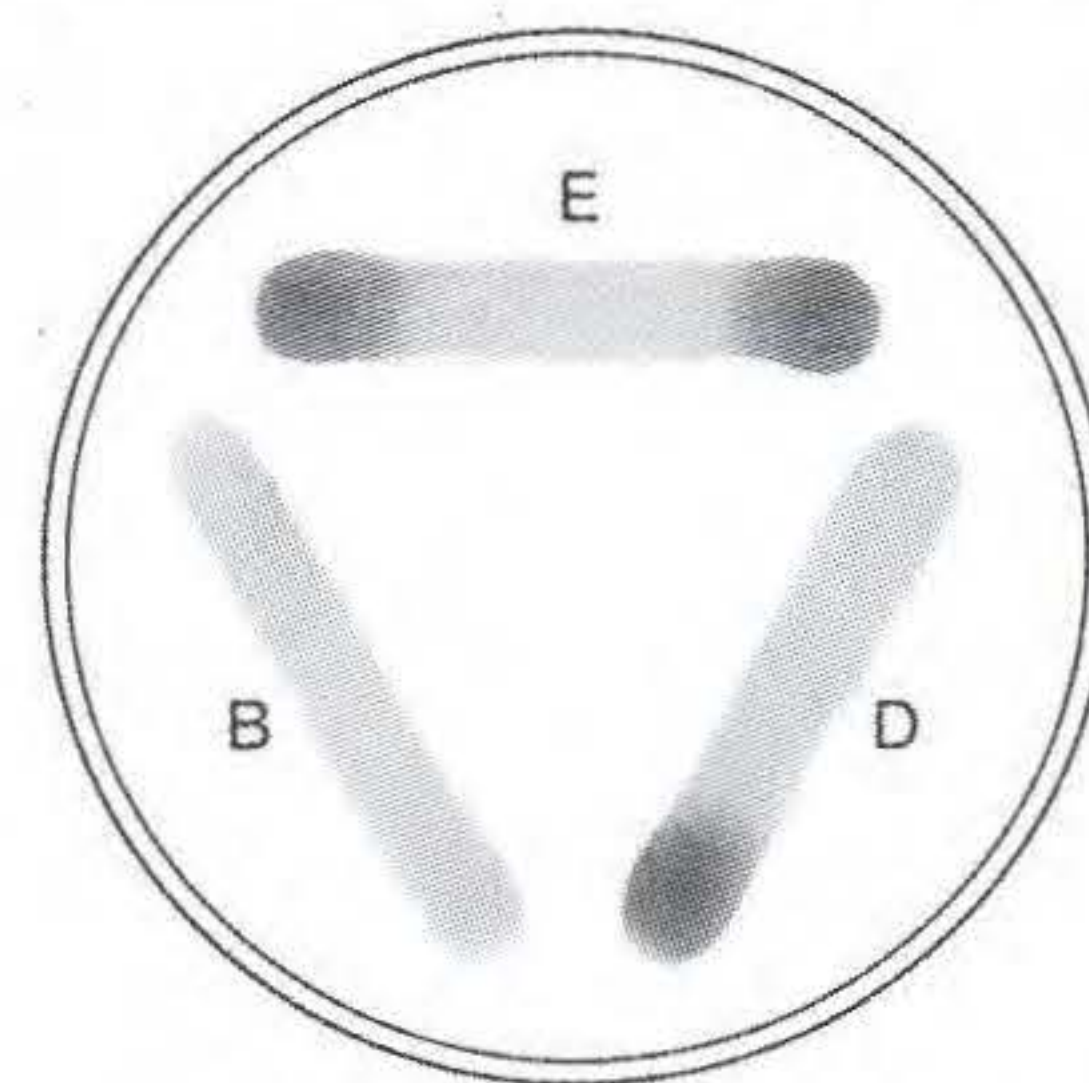
88	<i>P/-</i> ; <i>Q/-</i>	25	<i>p/p</i> ; <i>Q/-</i>
32	<i>P/-</i> ; <i>q/q</i>	14	<i>p/p</i> ; <i>q/q</i>

Esses resultados apóiam a hipótese de que a planta original era *P/p* ; *Q/q*?

48. Uma planta de fenótipo 1 foi autofecundada, e, na prole, havia 100 plantas com fenótipo 1 e 60 plantas de um fenótipo 2 alternativo. Esses números são compatíveis com as proporções esperadas 9:7, 13:3 e 3:1? Formule uma hipótese genética com base em seus cálculos.
49. Quatro linhagens homozigotas mutantes recessivas de *Drosophila melanogaster* (rotuladas de 1 a 4) mostraram coordenação anormal da perna, que tornava seu caminhar muito errático. Essas linhagens foram entrecruzadas; os fenótipos das moscas F₁ são mostrados no quadro a seguir, no qual “+” representa o andar do tipo selvagem e “-” representa o andar anormal.

	1	2	3	4
1	-	+	+	+
2	+	-	-	+
3	+	-	-	+
4	+	+	+	-

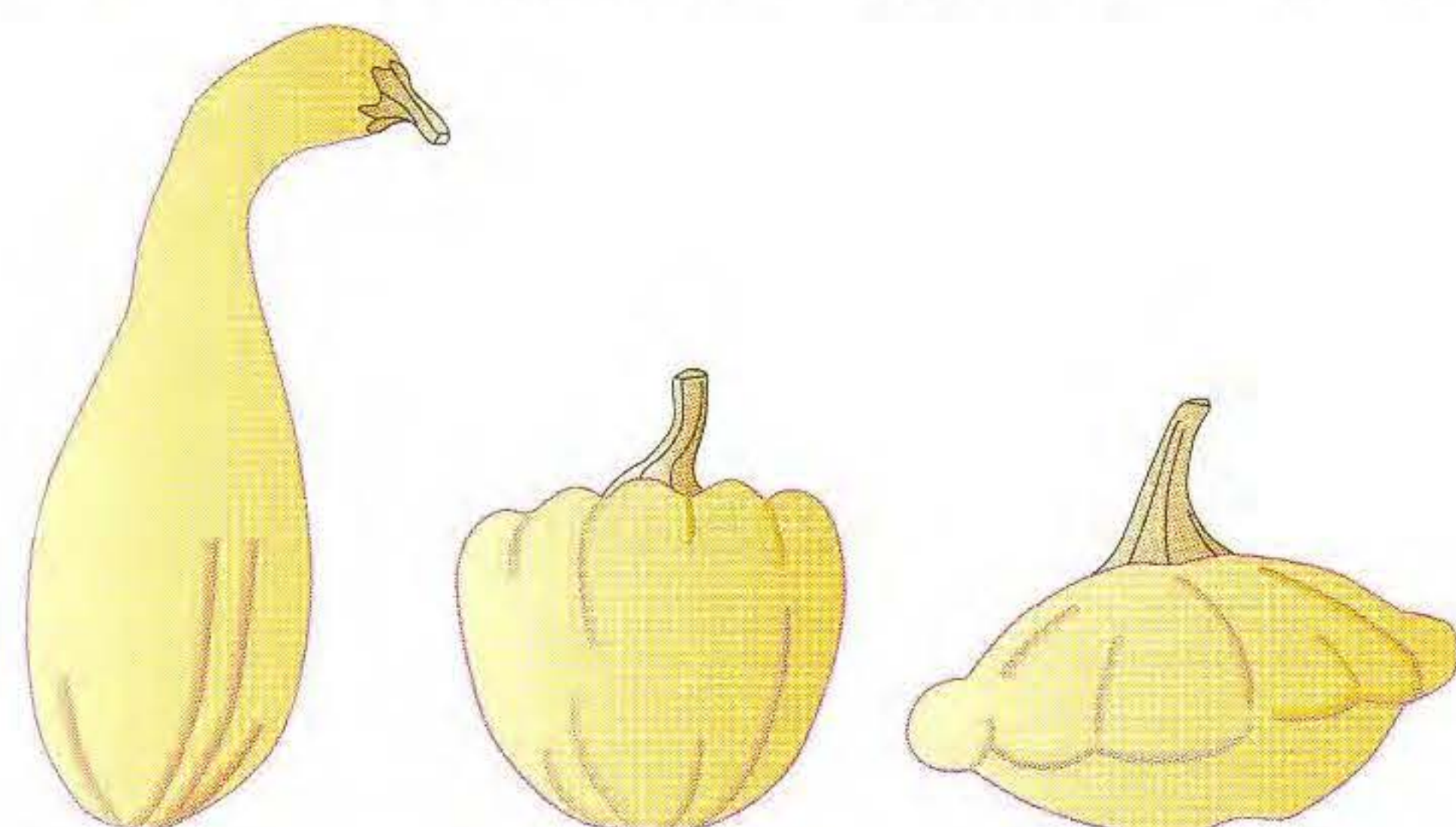
- a. Que tipo de teste essa análise representa?
- b. Quantos genes diferentes foram mutados na criação dessas quatro linhagens?
- c. Invente símbolos para tipo selvagem e mutantes e escreva os genótipos completos para todas as quatro linhagens e das moscas de F₁.
- d. Esses dados nos dizem quais genes estão ligados? Caso não, como a ligação pode ser testada?
- e. Esses dados nos dizem o número total de genes que tomam parte da coordenação da perna nesse animal?
50. Três mutantes independentemente isolados que requerem triptofano são chamados de *trpB*, *trpD* e *trpE*. As suspensões celulares de cada um são espalhadas em uma placa de meio nutritivo suplementado com triptofano apenas suficiente para permitir um fraco crescimento de uma linhagem *trp*. As faixas são dispostas em um padrão triangular, de modo que *não se tocam*. Um grande crescimento é notado em ambas as extremidades da faixa *trpE* e em uma ponta da faixa *trpD* (veja figura a seguir).



- a. Você acha que a complementação tem um papel?
- b. Explique resumidamente o padrão do intenso crescimento.
- c. Em que ordem na via de síntese de triptofano estão as etapas enzimáticas que são defectivas nos mutantes *trpB*, *trpD* e *trpE*?
- d. Por que é necessário adicionar uma pequena quantidade de triptofano ao meio para demonstrar tal padrão de crescimento?

PROBLEMAS DESAFIADORES

51. Uma linhagem pura de abóbora que produz frutos em forma de disco (veja a ilustração acompanhante) foi cruzada com uma linhagem pura que tem frutos longos. A F_1 tinha frutos discóides, mas a F_2 apresentou um novo fenótipo, esférico, e era composta das seguintes proporções:

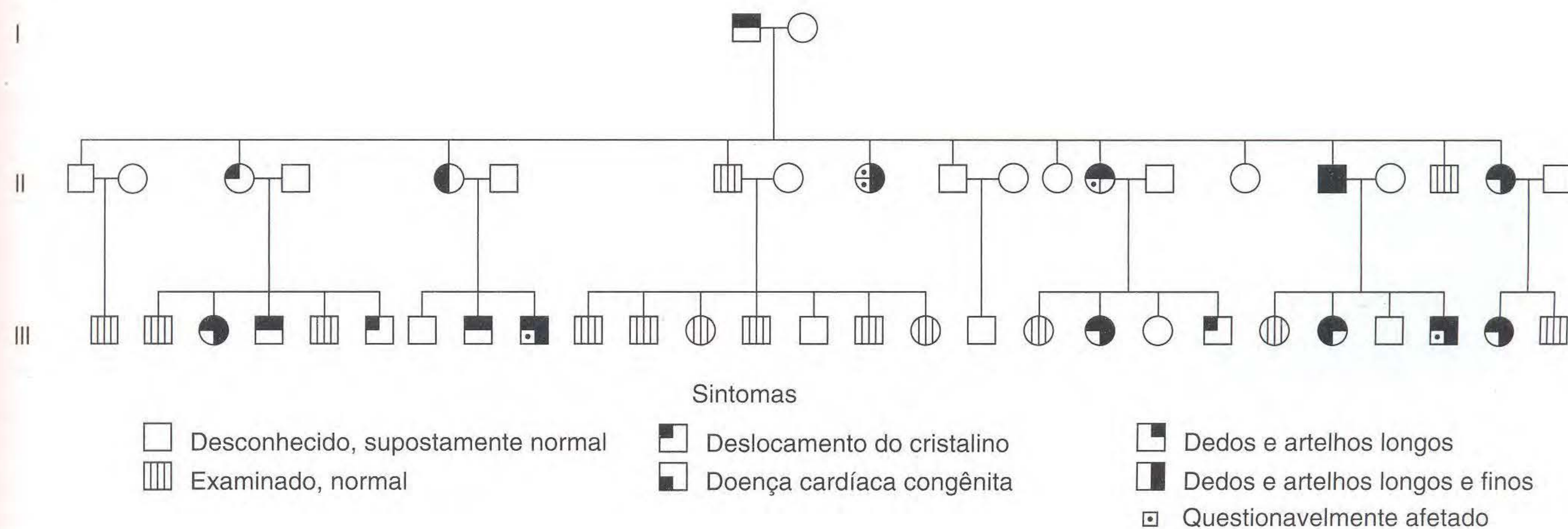


Longa

Esférica

Discóide

longa 32 esférica 178 discóide 270

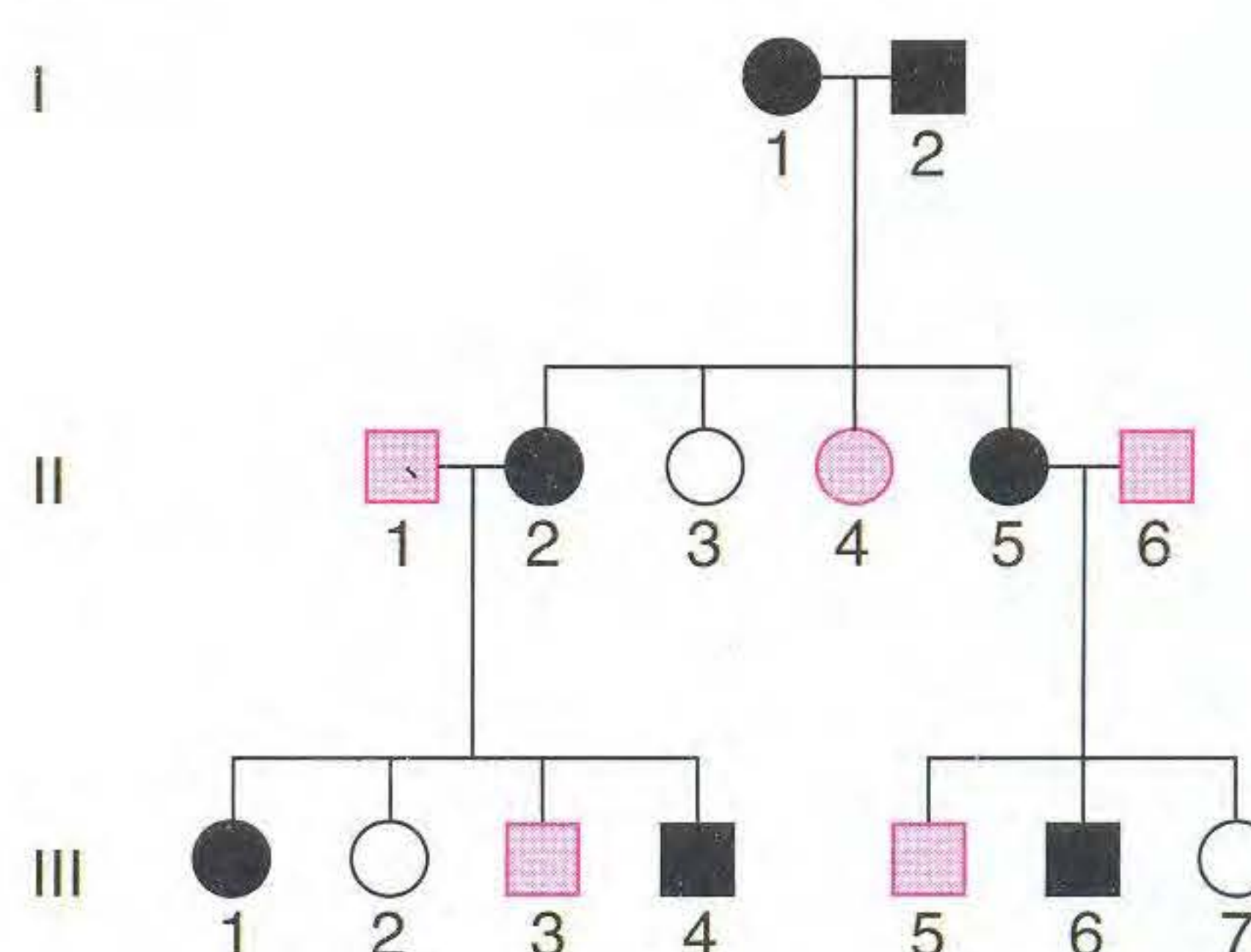


53. Em milho, três alelos dominantes, chamados A , C e R , devem estar presentes para produzir sementes coloridas. O genótipo $A/-$; $C/-$; $R/-$ é colorido; todos os outros são incolores. Uma planta colorida é cruzada com três plantas testadoras de genótipo conhecido. Com o testador a/a ; c/c ; R/R , a planta colorida produz 50% de sementes coloridas; com a/a ; C/C ; r/r , ela produz 25% coloridas; e com A/A ; c/c ; r/r , ela produz 50% coloridas. Qual o genótipo da planta colorida?
54. A produção de pigmento na camada externa das sementes de milho requer que cada um de três genes de segregação independente A , C e R esteja representado por, pelo menos, um alelo dominante, como especificado no Problema 53. O alelo dominante Pr de um quarto gene de segregação independente é necessário para converter o precursor bioquímico em um pigmento púrpura, e seu alelo recessivo pr torna o pigmento vermelho. As plantas que não produzem pigmento têm sementes amarelas. Considere um cruzamento de uma linhagem com o genótipo A/A ; C/C ; R/R ; pr/pr com uma linhagem do genótipo a/a ; c/c ; r/r ; Pr/Pr .
- Quais são os fenótipos dos genitores?
 - Qual será o fenótipo da F_1 ?

Proponha uma explicação para esses resultados e mostre os genótipos das gerações P , F_1 e F_2 . (Ilustração de P.J. Russell, *Genetics*, 3rd ed. HarperCollins, 1992.)

52. A síndrome de Marfan é um distúrbio do tecido conjuntivo fibroso, sendo caracterizada por muitos sintomas, incluindo dedos finos; defeitos oculares; doença cardíaca; e membros longos. (Flo Hyman, uma estrela do vôlei americano, sofria da síndrome de Marfan. Ela morreu de uma ruptura na aorta.)
- Use o heredograma a seguir para propor um modo de herança para a síndrome de Marfan.
 - Que fenômeno genético é mostrado por esse heredograma?
 - Especule um motivo para tal fenômeno.
- (Ilustração de J.V. Neel and W.J. Schull, *Human Heredity*. University of Chicago Press.)

- Que fenótipos, e em que proporções, aparecerão na prole de uma F_1 autofecundada?
 - Que proporções da prole você prevê do cruzamento-teste de uma F_1 ?
55. O alelo B dá aos camundongos uma pelagem preta, e b dá uma marrom. O genótipo e/e de outro gene de segregação independente impede a expressão de B e b , tornando a cor da pelagem bege, enquanto $E/-$ permite a expressão de B e b . Ambos os genes são autossômicos. No pedigree anexo, os símbolos pretos indicam pelagem preta, os símbolos rosa indicam marrom e os símbolos brancos indicam bege.



60. No trigo comum, *Triticum aestivum*, a cor dos grãos é determinada multiplicando-se genes duplicados, cada um com um alelo *R* e um *r*. Qualquer número de alelos *R* dará vermelho, e uma falta completa de alelos *R* dará o fenótipo branco. Em um cruzamento entre uma linhagem pura vermelha e uma pura branca, a F_2 era $\frac{63}{64}$ vermelha e $\frac{1}{64}$ branca.

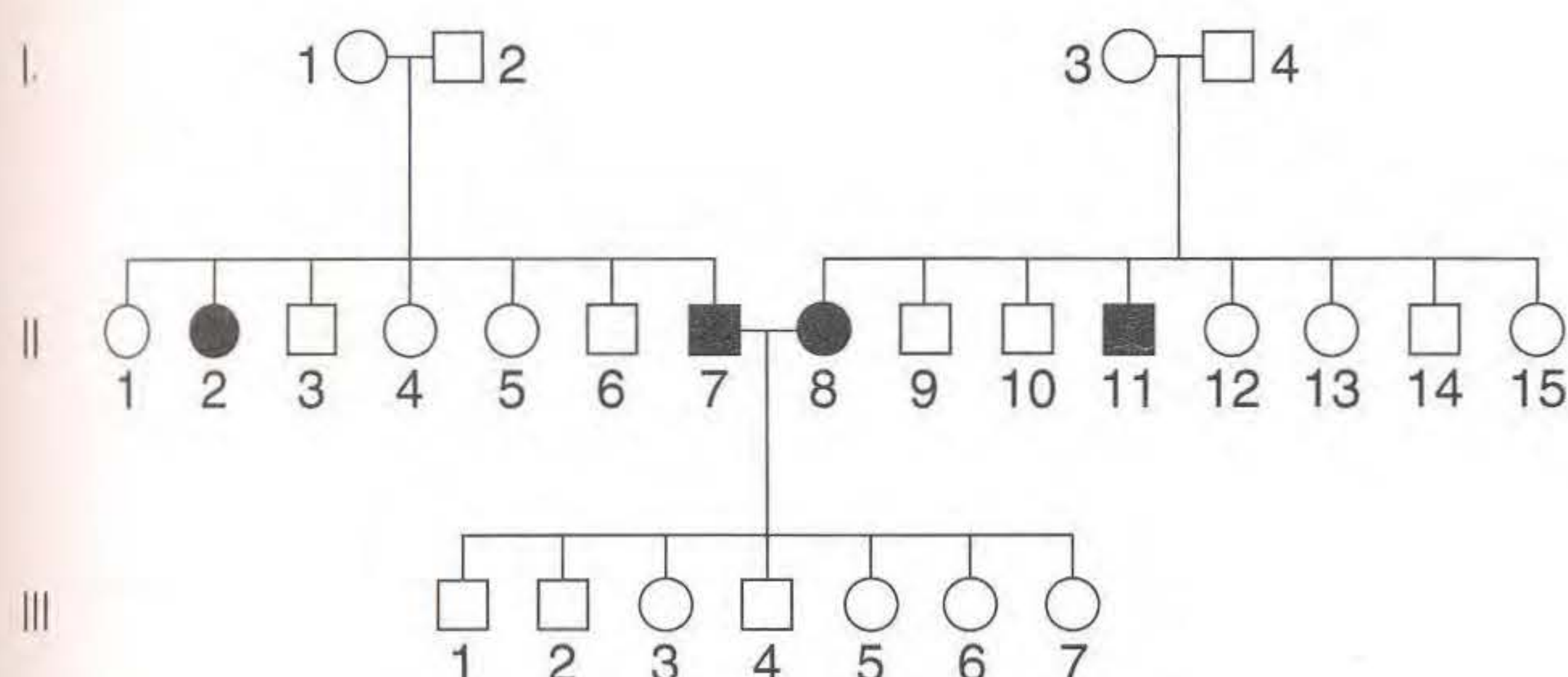
a. Quantos genes *R* estão segregando nesse sistema?

b. Mostre os genótipos dos genitores, a F_1 e a F_2 .

c. Diferentes plantas F_2 são retrocruzadas com o genitor branco. Cite exemplos de genótipos que dariam as seguintes proporções de prole em tais retrocruzamentos: (1) 1 vermelha : 1 branca, (2) 3 vermelhas : 1 branca, (3) 7 vermelhas : 1 branca.

d. Qual a fórmula que geralmente relaciona o número de genes segregantes à proporção de indivíduos vermelhos na F_2 em tais sistemas?

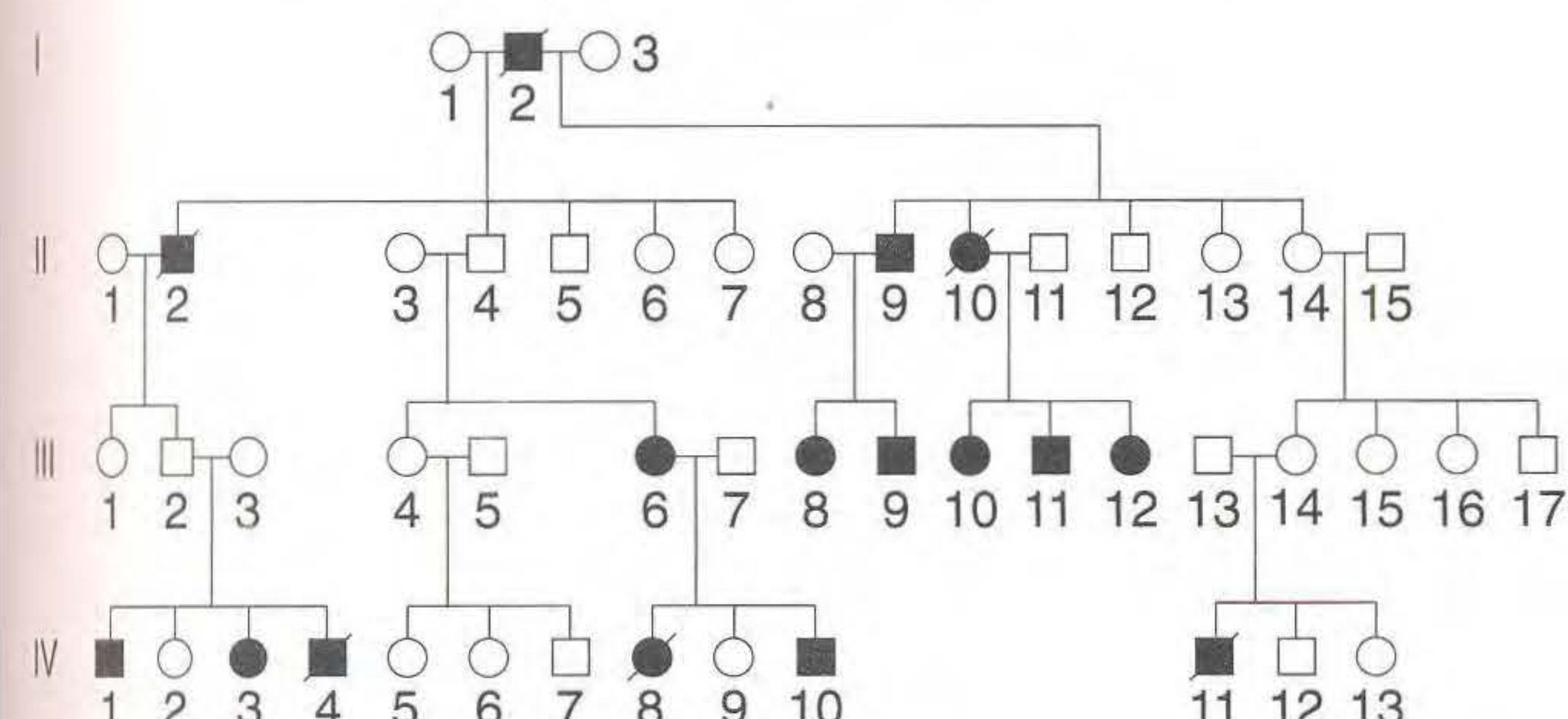
61. O heredograma seguinte mostra a herança de surdo-mutismo.



a. Forneça uma explicação para a herança dessa rara condição nas duas famílias nas gerações I e II, mostrando os genótipos de quantas pessoas possível; use símbolos de sua própria escolha.

b. Forneça uma explicação para a produção apenas de pessoas normais na geração III, certificando-se de que sua explicação é compatível com a resposta à parte a.

62. O heredograma a seguir é escleróticas azuis (fina parede externa do olho azulada) e ossos quebradiços.



●, ■ = Escleróticas azuis / = Ossos quebradiços

a. Essas duas anomalias são causadas pelo mesmo gene ou por genes separados? Descreva seu raciocínio claramente.

b. O gene (ou genes) são autossômicos ou ligados ao sexo?

c. O heredograma mostra alguma evidência de penetrância incompleta ou expressividade? Caso sim, faça os melhores cálculos que puder dessas medidas.

63. As operárias de uma linhagem de abelhas conhecidas como *Brown* (nada a ver com cor) apresentam o que é conhecido como “comportamento higiênico”; isto é, elas retiram o alimento das celas da colméia contendo pupas mortas e as removem da colméia. Esse comportamento impede a dispersão de bactérias infecciosas pela colônia. As operárias da linhagem *Van Scoy*, entretanto, não fazem essas ações, e, portanto, essa linhagem é chamada de “não-higiênica”. Quando uma rainha da linhagem *Brown* foi cruzada com machos *Van Scoy*, toda a F_1 era não-higiênica. Quando os machos dessa F_1 inseminaram uma rainha da linhagem *Brown*, os comportamentos da prole foram os seguintes:

$\frac{1}{4}$ de abelhas higiênicas

$\frac{1}{4}$ de abelhas que removiam o alimento mas não tiravam as pupas

$\frac{1}{2}$ de abelhas não-higiênicas

Entretanto, quando o compartimento com pupas mortas era aberto pelas operárias e as abelhas não-higiênicas eram depois examinadas, cerca de metade delas removiam as pupas mortas, mas a outra metade não.

a. Proponha uma hipótese genética para explicar esses padrões comportamentais.

b. Discuta os dados em relação à epistasia, dominância e interação ambiental.

(Nota: As operárias são estéreis, e todas as abelhas de uma linhagem levam os mesmos alelos.)

64. A cor normal de boca-de-leão é vermelha. Algumas linhagens puras mostrando variações de cor das flores foram encontradas. Quando essas linhagens puras foram cruzadas, elas deram os seguintes resultados (veja o quadro anexo):

Cruzamento	Genitores	F_1	F_2
1	laranja × amarelo	laranja	3 laranja : 1 amarelo
2	vermelho × laranja	vermelho	3 vermelho : 1 laranja
3	vermelho × amarelo	vermelho	3 vermelho : 1 amarelo
4	vermelho × branco	vermelho	3 vermelho : 1 branco
5	amarelo × branco	vermelho	9 vermelho : 3 amarelo : 4 branco
6	laranja × branco	vermelho	9 vermelho : 3 laranja : 4 branco
7	vermelho × branco	vermelho	9 vermelho : 3 amarelo : 4 branco

a. Explique a herança dessas cores.

b. Escreva os genótipos dos genitores, da F_1 e da F_2 de cada cruzamento.

65. Considere os seguintes indivíduos da F_1 em espécies diferentes e as proporções de F_2 produzidas por autofecundação:

F ₁	Proporção fenotípica na F ₂			
1 creme	$\frac{12}{16}$ creme	$\frac{3}{16}$ preto	$\frac{1}{16}$ cinza	
2 laranja	$\frac{9}{16}$ laranja	$\frac{7}{16}$ amarelo		
3 preto	$\frac{13}{16}$ preto	$\frac{3}{16}$ branco		
4 vermelho	$\frac{9}{16}$ vermelho	$\frac{3}{16}$ vermelho-salpicado	$\frac{4}{16}$ pequenos pontos vermelhos	

Se cada F₁ for submetida a um cruzamento-teste, que proporções fenotípicas resultarão na prole do cruzamento-teste?

66. Para compreender a base genética da locomoção no nematódeo diplóide *Caenorhabditis elegans*, foram obtidas mutações recessivas, todas tornando o verme meneador inefetivo ao invés de se mover com seu deslizar suave. Essas mutações supostamente afetam os sistemas nervoso ou muscular. Foram entrecruzados 12 mutantes homozigotos, e os híbridos de F₁ foram examinados para ver se meneavam. Os resultados foram os seguintes, onde um sinal mais significa que o híbrido de F₁ era tipo selvagem (deslizavam) e “w” significa que o híbrido meneava.

	1	2	3	4	5	6	7	8	9	10	11	12
1	w	+	+	+	w	+	+	+	+	+	+	+
2		w	+	+	+	w	+	w	+	w	+	+
3			w	w	+	+	+	+	+	+	+	+
4				w	+	+	+	+	+	+	+	+
5					w	+	+	+	+	+	+	+
6						w	+	w	+	w	+	+
7							w	+	+	+	w	w
8								w	+	w	+	+
9									w	+	+	+
10										w	+	+
11											w	w
12												w

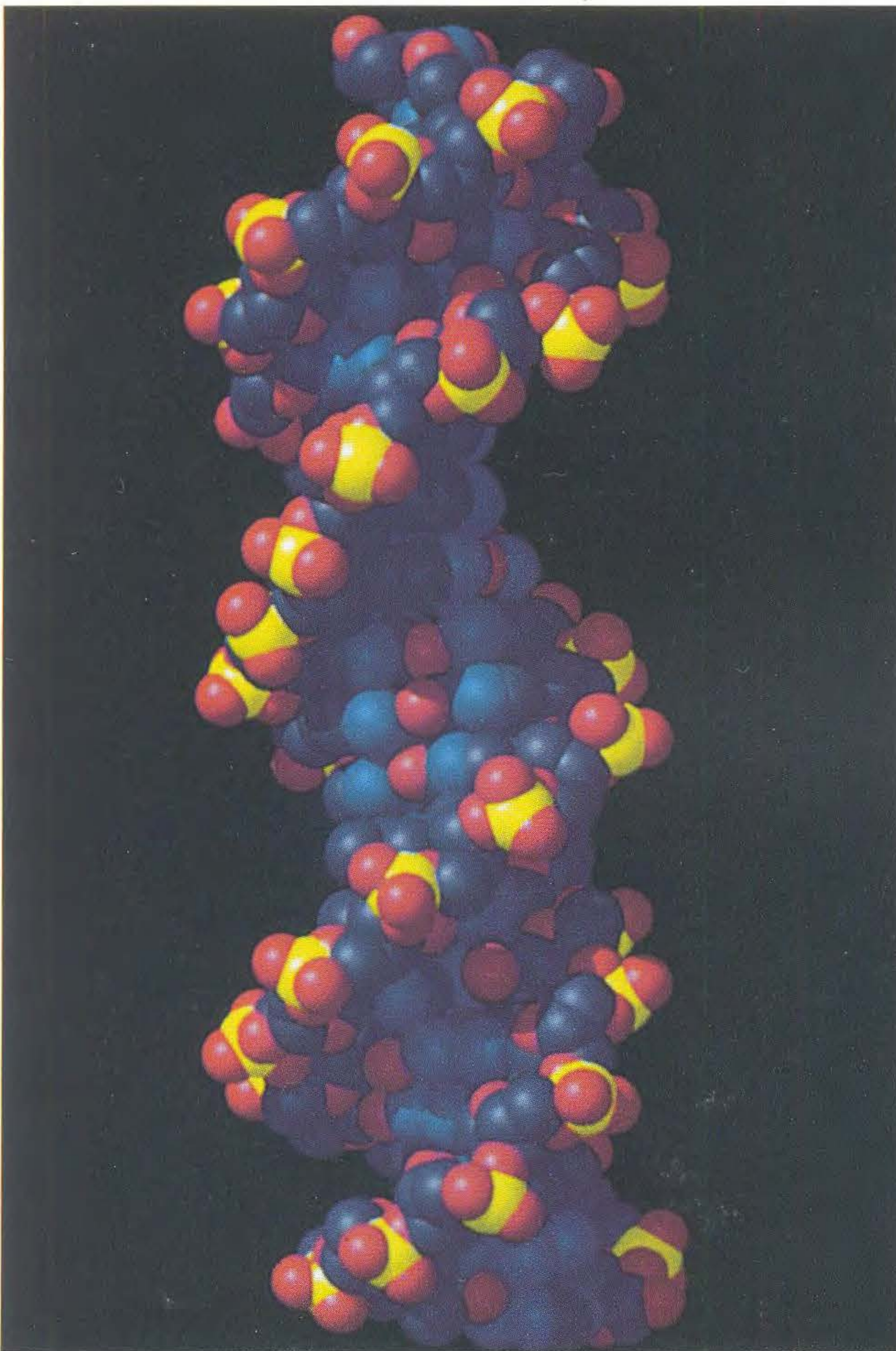
- a. Explique qual o objetivo desse teste.
- b. Use esse raciocínio para atribuir genótipos a todos os 12 mutantes.
- c. Explique por que o fenótipo dos híbridos de F₁ entre os mutantes 1 e 2 diferiu do dos híbridos entre os mutantes 1 e 5.
67. Um geneticista trabalhando com um fungo haplóide fez um cruzamento entre dois mutantes de crescimento lento chamados *mossy* e *spider* (referindo-se ao aspecto anormal das colônias). As tétrades do cruzamento são de três tipos (A, B, C), mas dois deles continham esporos que não germinaram.

Esporo	A	B	C
1	tipo selvagem	tipo selvagem	spider
2	tipo selvagem	spider	spider
3	sem germinação	mossy	mossy
4	sem germinação	sem germinação	mossy

Crie um modelo para explicar esses resultados genéticos, e proponha uma base molecular para seu modelo.

DNA: Estrutura e Replicação

7



Modelo computadorizado do DNA. [J. Newdol, Computer Graphics Laboratory, University of California, San Francisco. Copyright by Regents, University of California.]

James Watson (um geneticista microbiano americano) e Francis Crick (um físico inglês) descobriram a estrutura do DNA em 1953. Seu modelo da estrutura do DNA foi revolucionário. Eles propuseram uma definição de gene em termos químicos e, fazendo isso, abriram o caminho para a compreensão da ação gênica e da hereditariedade a nível molecular. Uma medida da importância de sua descoberta é que a estrutura em dupla hélice tornou-se um ícone que é cada vez mais visto frequentemente em pinturas, em esculturas e, mesmo, em *playgrounds* (Figura 7.1).

A história começa na primeira metade do século XX, quando o resultado de vários experimentos levou os cientistas a concluir que o DNA é o

Perguntas Fundamentais

- Antes da descoberta da dupla hélice, qual a evidência experimental de que o DNA é o material genético?
- Que dados foram usados para deduzir o modelo da dupla hélice do DNA?
- Como a estrutura da dupla hélice sugere um mecanismo para a replicação do DNA?
- Como a replicação do DNA pode ser tanto rápida quanto precisa?
- Que mecanismo especial replica as pontas dos cromossomos, e quais são as consequências para a saúde humana se o final da replicação for defeituoso?

Tópicos

- 7.1 DNA: o material genético
- 7.2 A estrutura do DNA
- 7.3 Replicação semiconservativa
- 7.4 Visão geral da replicação do DNA
- 7.5 O replissomo: uma marcante máquina de replicação
- 7.6 Replicação em organismos eucarióticos
- 7.7 Telômeros e telomerase: término da replicação

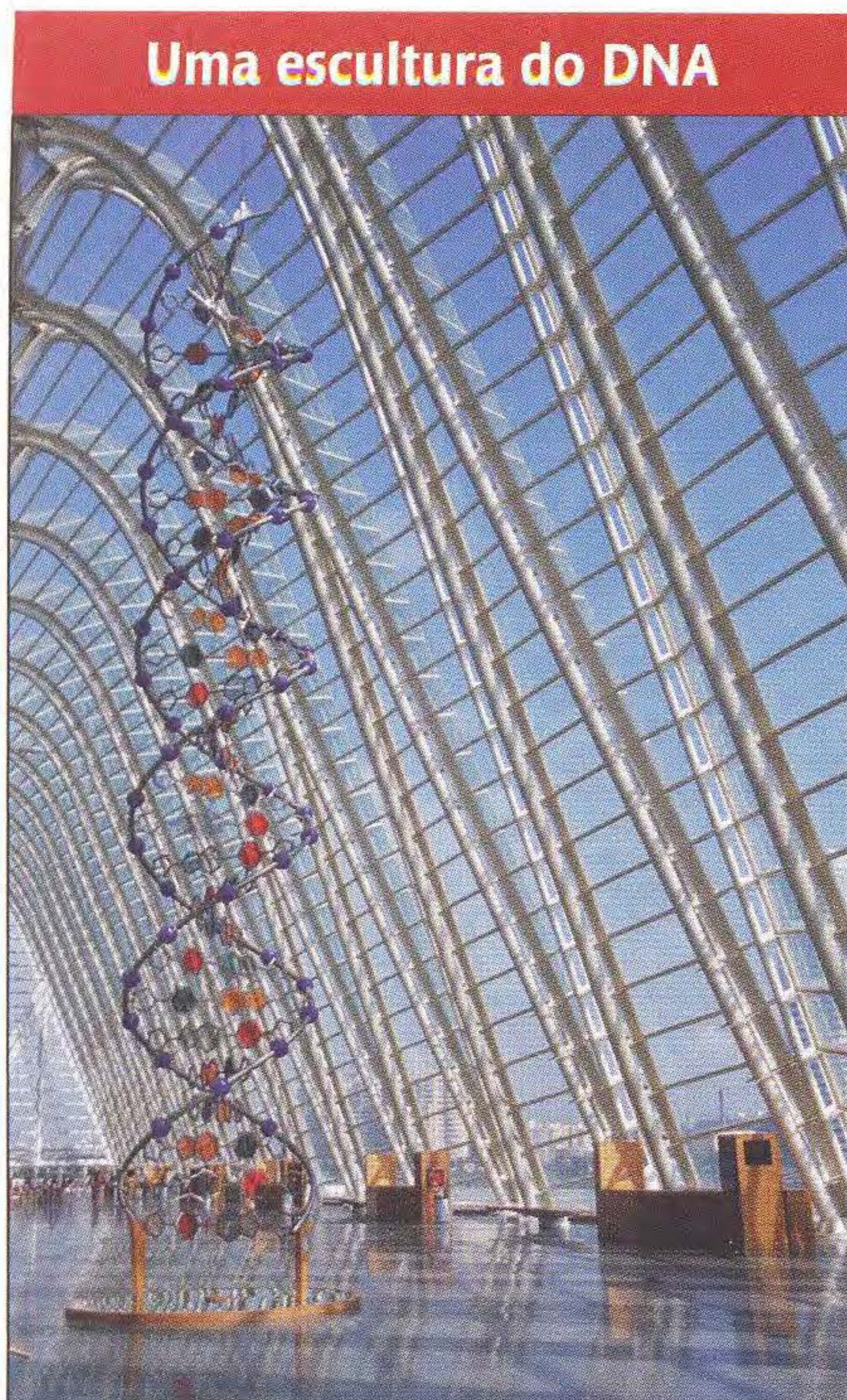


FIG. 7.1 Esse modelo de Watson-Crick da molécula de DNA está no Museum of Sciences, Valencia, Espanha. [Arco Images/Alamy.]

material genético, e não outra molécula biológica tal como um carboidrato, proteína ou lipídeo. O DNA é uma molécula simples feita apenas de quatro blocos estruturais (os quatro nucleotídeos). Assim, foi necessário compreender como essa molécula tão simples podia ser o mapa para a incrível diversidade de organismos na Terra.

O modelo da dupla hélice proposto por Watson e Crick foi feito com base nos resultados de cientistas anteriores a eles. Eles se basearam em descobertas anteriores da composição química do DNA e proporções de suas bases. Além disso, imagens de difração de raios X revelaram ao olho treinado que o DNA é uma hélice de dimensões precisas. Watson e Crick concluíram que o DNA é uma dupla hélice composta de duas cromátides de nucleotídeos ligados que se enrolam uma na outra.

A estrutura proposta do material hereditário imediatamente sugeriu como ele poderia servir como um mapa, e como esse mapa poderia ser transmitido ao longo das gerações. Primeiro, a informação para fazer um organismo está codificada na sequência de bases dos nucleotídeos que compõem os dois filamentos de DNA da hélice. Segundo, devido às regras de complementariedade de bases descobertas por Watson e Crick, a sequência de um filamento determina a sequência do outro filamento. Desse modo, a informação genética na sequência de DNA pode ser transmitida de uma geração para a seguinte, cada um dos filamentos separados de DNA servindo como um molde para produzir novas cópias da molécula.

Neste capítulo, enfocaremos o DNA, sua estrutura e a produção de cópias do DNA em um processo chamado de replicação. Exatamente como o DNA é replicado ainda é uma área ativa de pesquisa 50 anos após a descoberta da dupla hélice.

Nossa compreensão atual do mecanismo de replicação dá um papel central a uma máquina proteica, chamada replissomo. Esse complexo de proteínas coordena as numerosas reações que são necessárias para a rápida e precisa replicação do DNA.

► 7.1 DNA: o Material Genético

Antes de vermos como Watson e Crick resolveram a estrutura do DNA, vamos revisar o que se sabia sobre os genes e DNA na época em que eles começaram sua histórica colaboração:

1. Genes — os “fatores” hereditários descritos por Mendel — eram conhecidos como associados a características específicas, mas sua natureza física não era compreendida. Similarmente, as mutações eram conhecidas como alterando o funcionamento do gene, mas não era compreendido o que uma mutação era exatamente.
2. A hipótese um gene — uma proteína (descrita no Capítulo 6) postula que os genes controlam a estrutura das proteínas.
3. Os genes são conhecidos como levados nos cromossomos.
4. Os cromossomos consistem em DNA e proteínas.
5. Os resultados de uma série de experimentos começando nos anos de 1920 revelaram que o DNA era o material genético. Esses experimentos, descritos em seguida, mostraram que as células bacterianas que expressam um fenótipo podem ser transformadas em células que expressam um fenótipo diferente, e que o agente transformante é o DNA.

Descoberta da transformação

Frederick Griffith fez uma curiosa observação no curso de experimentos na bactéria *Streptococcus pneumoniae* realizados em 1928. Esta bactéria, que causa pneumonia em humanos, é normalmente letal em camundongos. Entretanto, algumas linhagens dessa espécie bacteriana evoluíram para serem menos virulentas (menos capazes de causar doença ou morte). Os experimentos de Griffith estão resumidos na Figura 7.2. Nesses experimentos, Griffith usou duas linhagens que são distinguíveis pelo aparecimento de suas colônias quando cultivadas em culturas de laboratórios. Uma linhagem era um tipo virulento mortal para a maioria dos animais de laboratório. As células dessa linhagem são encapsuladas em polissacarídeo, dando às colônias um aspecto liso; assim, essa linhagem é identificada como S. A outra linhagem de Griffith era um tipo não-virulento mutante que cresce em camundongos, mas não é letal. Nessa linhagem, a capa de polissacarídeo está ausente, dando colônias de aspecto rugoso; essa linhagem é chamada R.

Griffith matou algumas células virulentas por aquecimento. Ele então injetou as células mortas por calor em camundongos. Os camundongos sobreviveram, mostrando que o envoltório das células não causa a morte. Entretanto, os camundongos injetados com uma mistura de células viru-

Transformação das células R em células S

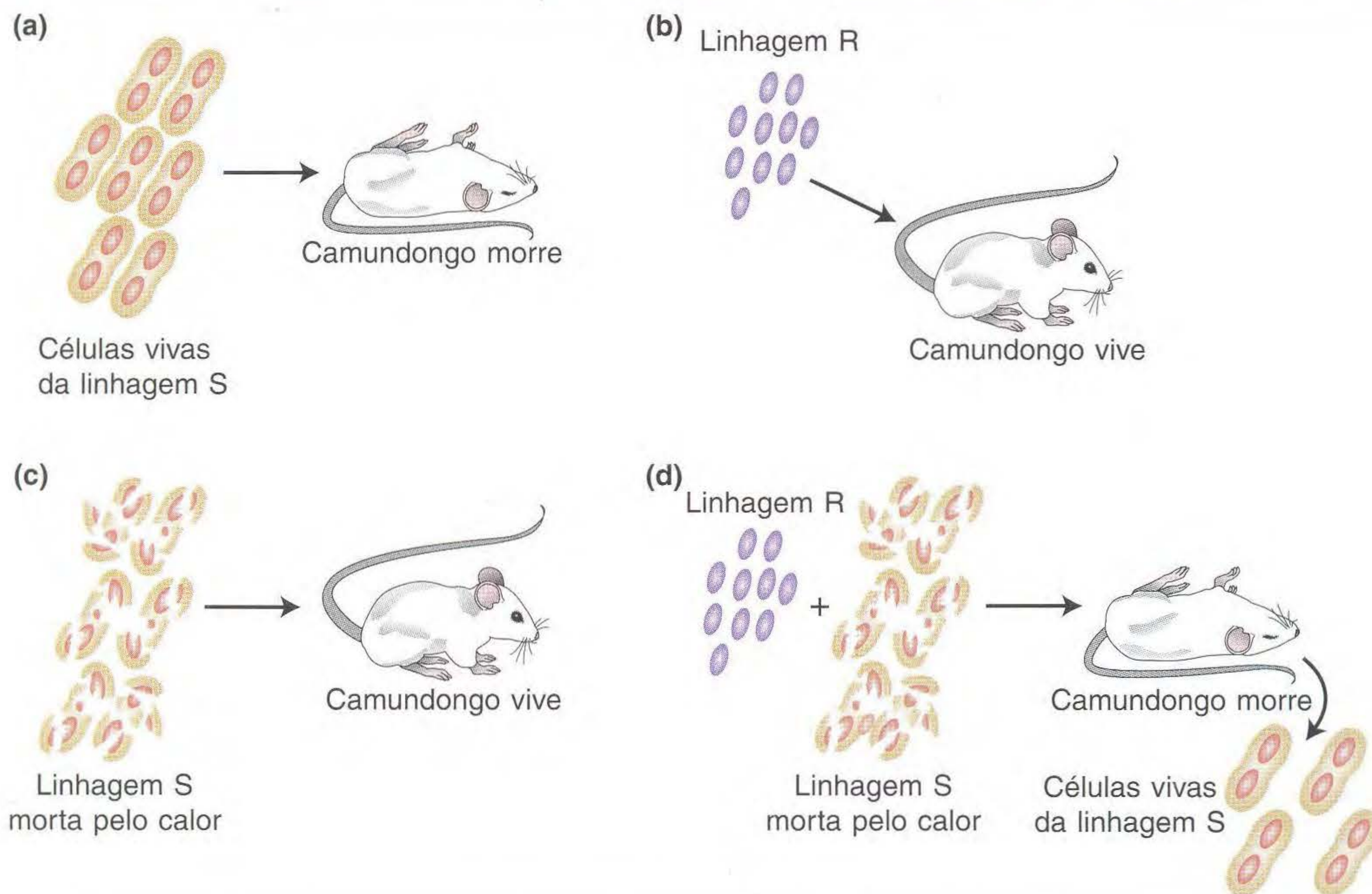


FIG. 7.2 A presença de células S mortas por aquecimento transforma as células R vivas em células S vivas. (a) O camundongo morre após a injeção da linhagem S virulenta. (b) O camundongo sobrevive após a injeção da linhagem R. (c) O camundongo sobrevive após a injeção de linhagem S morta por aquecimento. (d) O camundongo morre após a injeção com uma mistura de células mortas por aquecimento e linhagem R viva. A linhagem S morta por aquecimento de algum modo transforma a linhagem R em virulência. [G.S. Stent and R. Calendar, *Molecular Genetics*, 2nd ed. Copyright 1978 by W.H. Freeman and Company. R. Sager and F. J. Ryan, *Cell Heredity*, Wiley, 1961.]

lentas mortas pelo calor e células vivas não-virulentas morreram. Além disso, as células vivas podiam ser recuperadas dos camundongos mortos. Essas células davam colônias lisas e eram virulentas em uma subsequente injeção. De algum modo, os restos celulares das células S aquecidas tinham convertido células R vivas em células S vivas. O processo, já discutido no Capítulo 5, é chamado de *transformação*.

A etapa seguinte foi determinar que componente químico das células doadoras mortas tinha causado essa transformação. Essa substância tinha mudado o genótipo da linhagem receptora e, portanto, podia ser candidata a ser o material genético. Esse problema foi resolvido por experimentos feitos em 1944 por Oswald Avery com a colaboração de Colin MacLeod e Mcllyn MacCarty (Figura 7.3). Seu enfoque ao problema foi destruir quimicamente todas as principais categorias de substâncias químicas no extrato das células mortas, uma de cada vez, e descobrir se o extrato tinha perdido a habilidade da transformação. As células virulentas tinham um revestimento polissacarídico liso, enquanto as células virulentas não. Assim, os polissacarídeos eram um candidato óbvio a ser o agente transformante. Entretanto, quando os polissacarídeos foram destruídos, a mistura ainda podia transformar. Proteínas, gorduras e ácidos ribonucleicos (RNA) foram todos mostrados como não sendo o agente transformante. A mistura perdia sua habilidade transformante apenas quando a mistura doadora era tratada com a enzima desoxirribonuclease (DNase), que quebra o DNA. Esses resultados indicaram fortemente que o DNA era o material genético. Hoje é sabido

que os fragmentos do DNA transformante que conferem a virulência entram no cromossomo bacteriano e substituem suas contrapartes que conferem não-virulência.

A demonstração de que o DNA é o princípio transformante foi a primeira demonstração de que os genes (o material hereditário) são compostos de DNA.

Experimento de Hershey-Chase

Os experimentos conduzidos por Avery *et al.* foram definitivos, mas muitos cientistas relutaram em aceitar o DNA (e não as proteínas) como o material genético. Como uma molécula de tão baixa complexidade como o DNA podia codificar a diversidade de vida deste planeta? Alfred Hershey e Martha Chase deram evidências adicionais, em 1952, em um experimento que usou o fago T2, um vírus que infecta bactérias. Eles raciocinaram que o fago infectante devia injetar na bactéria a informação específica que dita a reprodução de novas partículas virais. Se eles pudessem descobrir que material o fago estava injetando na bactéria hospedeira, eles teriam determinado o material genético dos fagos.

O fago tem constituição molecular relativamente simples. A maioria de sua estrutura é de proteínas, com o DNA contido dentro da bainha de proteínas de sua “cabeça”. Hershey e Chase decidiram marcar diferencialmente o DNA e a proteína usando radioisótopos, de modo que podiam rastrear os

DNA é o agente transformante

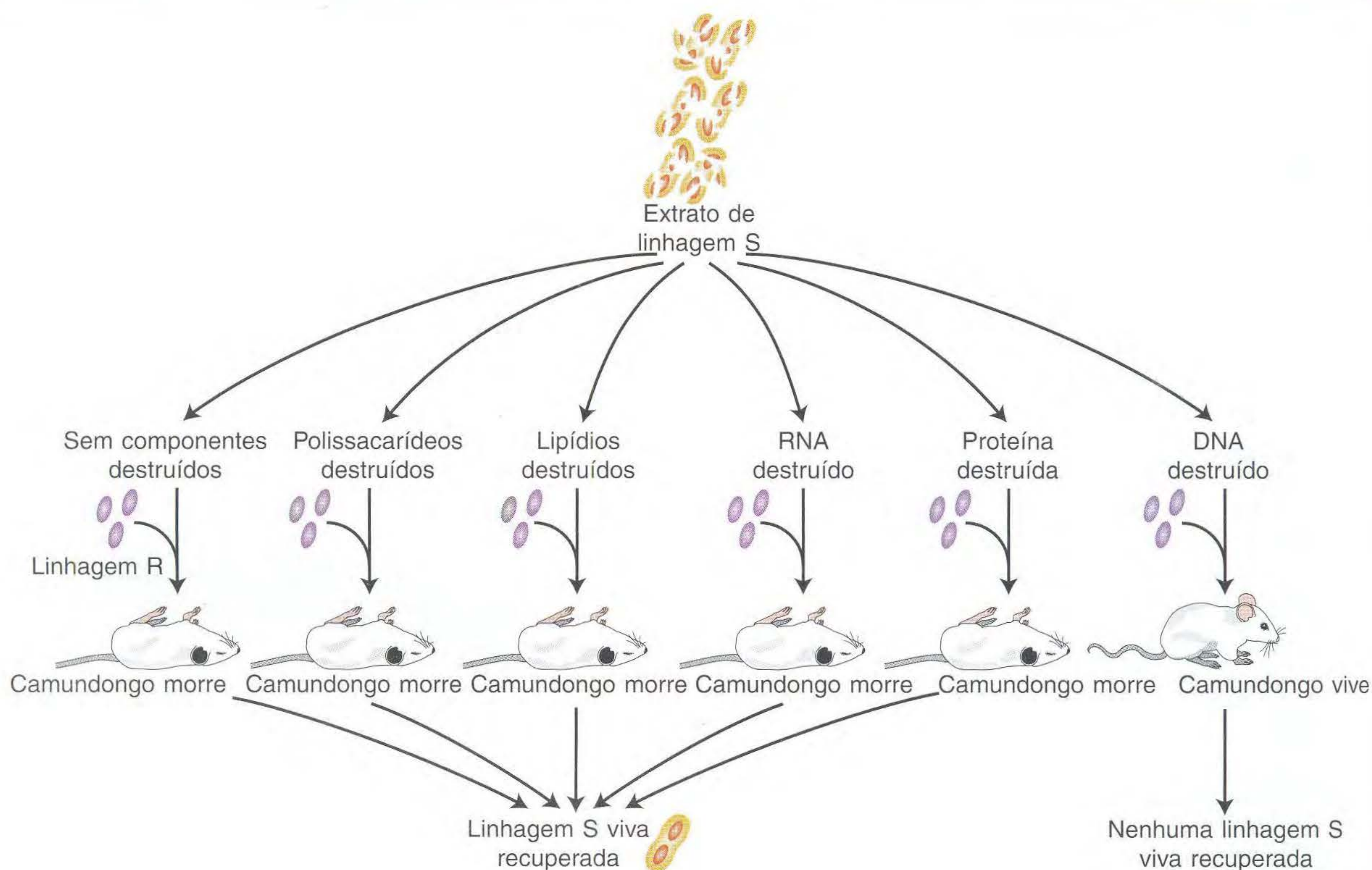


FIG. 7.3 O DNA é o agente transformante da linhagem R em virulência. Se o DNA em um extrato de células da linhagem S mortas por aquecimento é destruído, então os camundongos são injetados com uma mistura das células mortas por aquecimento e as células da linhagem R não-virulentas vivas não são mais mortas.

dois materiais durante a infecção. O fósforo não é encontrado nas proteínas, mas é parte integral do DNA. Contrariamente, o enxofre está presente nas proteínas, mas nunca no DNA. Hershey e Chase incorporaram o radioisótopo do fósforo (^{32}P) ao DNA do fago e o do enxofre (^{35}S) às proteínas de uma cultura separada de fagos. Como mostrado na Figura 7.4, eles então injetaram duas culturas de *E. coli* com muitas partículas de vírus por célula: uma cultura de *E. coli* recebeu fago marcado com ^{32}P e a outra recebeu fago marcado com ^{35}S . Após dar tempo suficiente para que ocorra a infecção, eles fragmentaram o fago vazio (chamados de *ghosts*) das células bacterianas por agitação em um liquidificador. Eles separaram as células bacterianas dos *ghosts* de fagos em uma centrífuga e, então, dosaram a radioatividade nas duas frações. Quando os fagos marcados com ^{32}P foram usados para infectar *E. coli*, a maior parte da radioatividade estava dentro das células bacterianas, indicando que o DNA do fago entrou nas células. Quando foram usados fagos marcados com ^{35}S , a maior parte do material radioativo terminou nos *ghosts* dos fagos, indicando que a proteína do fago nunca entra na célula bacteriana. A conclusão é inevitável: o DNA é o material hereditário. As proteínas do fago eram meras embalagens estruturais que são descartadas após a entrada do DNA viral na célula bacteriana.

► 7.2 A Estrutura do DNA

Mesmo antes de a estrutura do DNA ser elucidada, os estudos genéticos indicaram que o material genético deve ter três propriedades principais:

1. Como essencialmente cada célula do corpo de um organismo tem a mesma constituição genética, a replicação fiel do material genético em cada divisão celular é crucial. Assim, as características estruturais do DNA *devem permitir uma replicação fiel*. Essas características estruturais serão consideradas mais adiante neste capítulo.
2. Como ele deve codificar uma constelação de proteínas expressas por um organismo, o material genético *deve ter conteúdo informacional*. Como a informação codificada no DNA é decifrada para produzir proteínas será o objeto dos Capítulos 8 e 9.
3. Como as alterações hereditárias, chamadas mutações, fornecem a matéria-prima para a seleção evolutiva, o material genético *deve ser capaz de mudar* em raras ocasiões. Entretanto, a estrutura do DNA deve ser estável de modo que os organismos possam se basear em sua informação codificada. Consideraremos os mecanismos de mutação no Capítulo 15.

O material genético do fago é o DNA

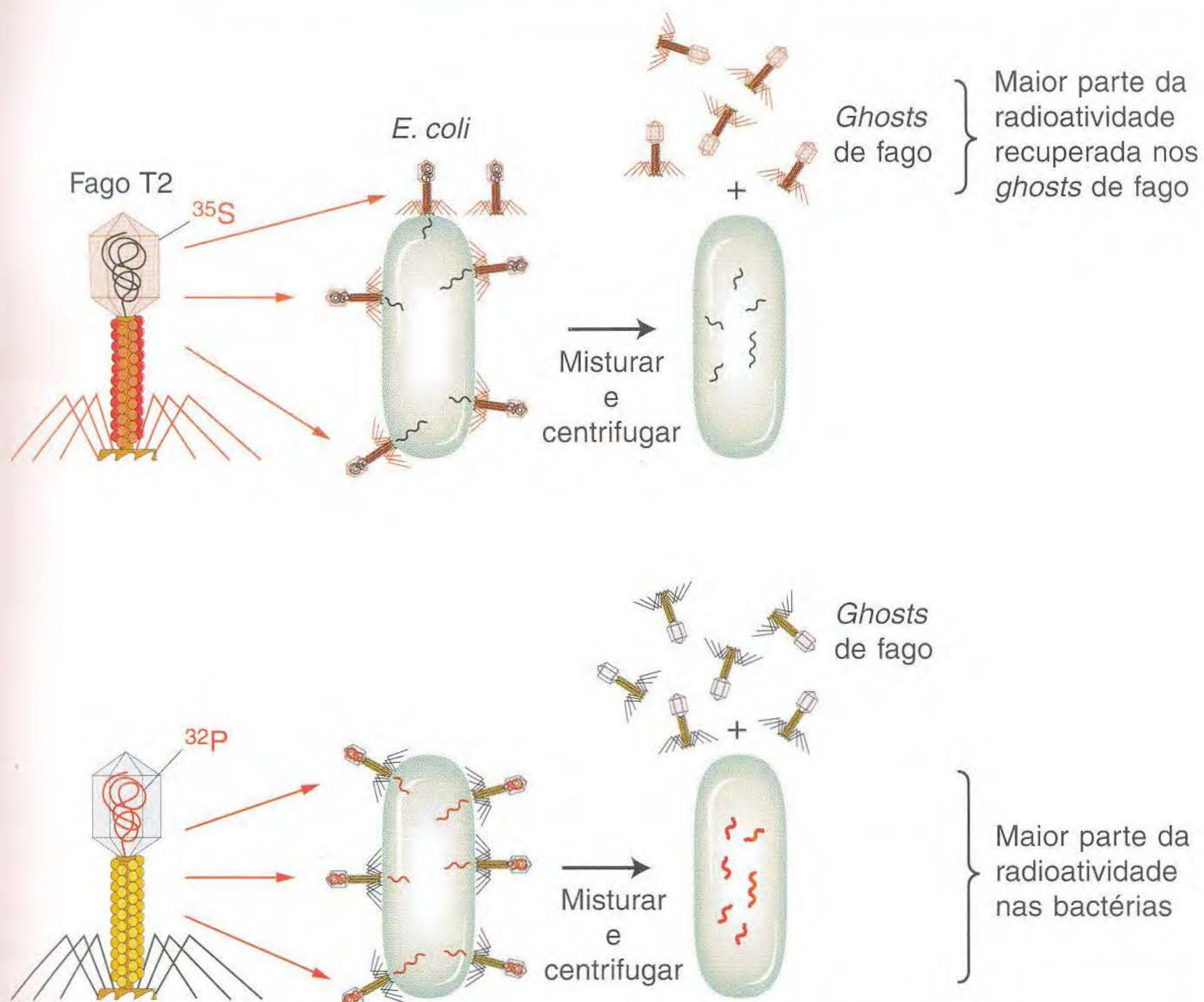


FIG. 7.4 O experimento de Hershey-Chase demonstrou que o material genético dos fagos é o DNA, não a proteína. O experimento usa dois conjuntos de bacteriófagos T2. Em um conjunto, a capa de proteína é marcada com enxofre radioativo (^{35}S), não encontrado no DNA. No outro conjunto, o DNA é marcado com fósforo radioativo (^{32}P), não encontrado na proteína. Apenas o ^{32}P é injetado na *E. coli*, indicando que o DNA é o agente necessário para a produção de novos fagos.

A estrutura do DNA antes de Watson e Crick

Considere a descoberta da estrutura da dupla hélice de DNA por Watson e Crick como a solução de um complicado quebra-cabeças tridimensional. Para resolvê-lo, Watson e Crick usaram um processo chamado “construção de modelo” no qual eles reuniram os resultados de experimentos anteriores e continuados (as peças do quebra-cabeças) para formar o quebra-cabeças tridimensional (o modelo da dupla hélice). Para compreender como o fizeram, primeiro precisamos conhecer que peças do quebra-cabeças estavam disponíveis a Watson e Crick em 1953.

Os blocos estruturais do DNA A primeira peça do quebra-cabeças era o conhecimento dos blocos estruturais básicos do DNA. Como substância química, o DNA é bem simples. Ele contém três tipos de componentes químicos: (1) **fosfato**, (2) um açúcar chamado **desoxirribose** e (3) quatro **bases** nitrogenadas: adenina, guanina, citosina e timina. Os átomos de carbono nas bases têm números para facilitar a referência. Os átomos de carbono no grupo açúcar também têm números. Nesse caso, o número é seguido de um primo (1', 2', e assim por diante). O açúcar no DNA é chamado de “desoxirribose” porque tem apenas um átomo de hidrogênio (H) no carbono 2', ao contrário da ribose (um componente do RNA), que tem um grupo hidroxila (OH) nessa posição. Duas das bases, adenina e guanina, têm uma estrutura de dois anéis característica de um tipo de substância chamada uma **purina**. As

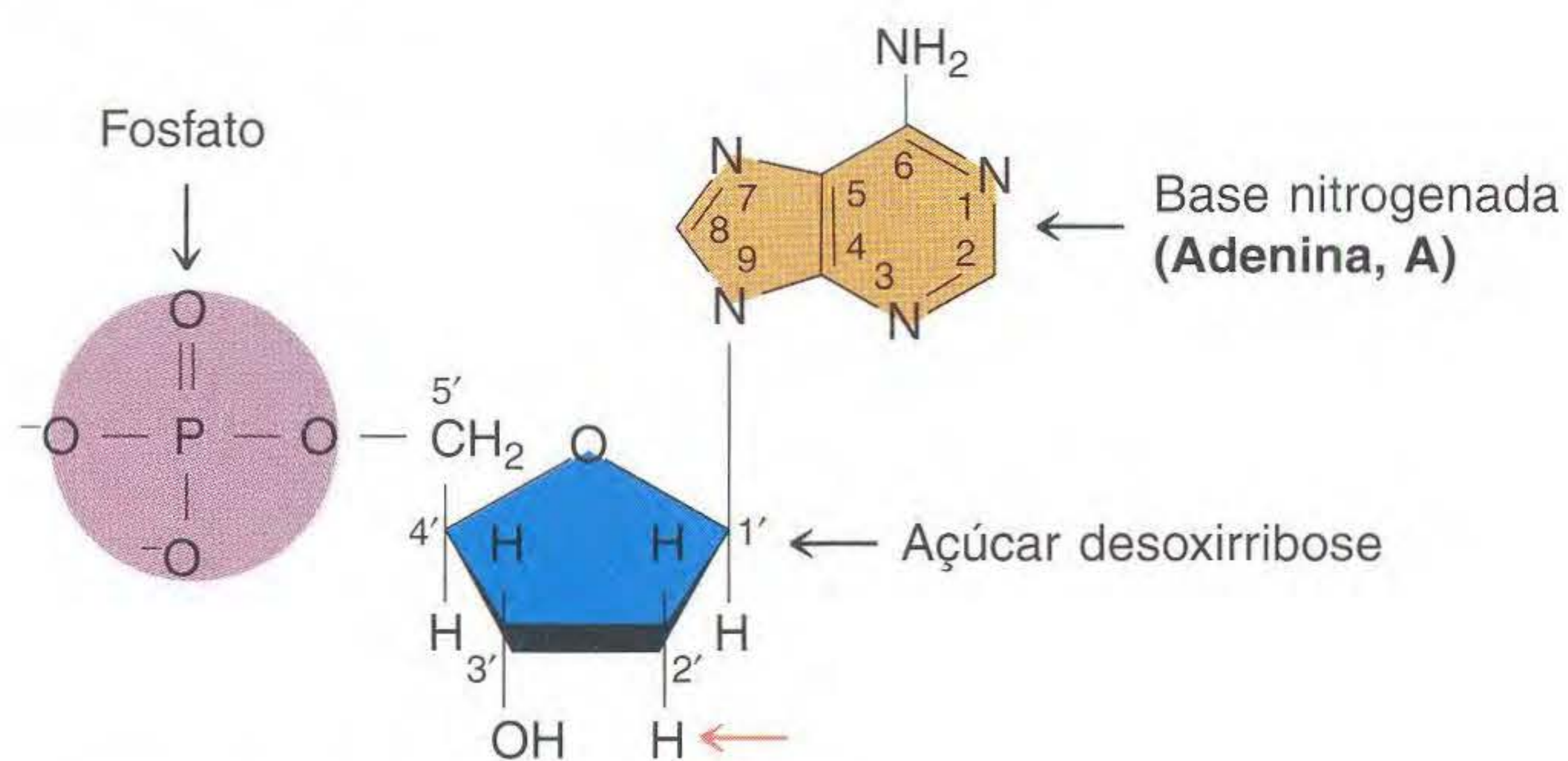
outras duas bases, citosina e timina, têm uma estrutura com um só anel, chamada **pirimidina**. Os componentes químicos do DNA são dispostos em grupos chamados **nucleotídeos**, cada um composto de um grupo fosfato, uma molécula de desoxirribose e qualquer uma das quatro bases (Figura 7.5). É conveniente chamar cada nucleotídeo pela primeira letra do nome de sua base: A, G, C ou T. O nucleotídeo com a base adenina é chamado de desoxiadenosina 5'-monofosfato, em que o 5' refere-se à posição do átomo de carbono no anel de açúcar ao qual o único grupo (mono) fosfato é ligado.

As regras de Chargaff da composição de bases A segunda peça do quebra-cabeças usado por Watson e Crick veio de um trabalho feito anos antes por Erwin Chargaff. Estudando uma grande seleção de DNA de organismos diferentes (Quadro 7.1), Chargaff estabeleceu algumas regras empíricas sobre as quantidades de cada tipo de nucleotídeo encontrado no DNA:

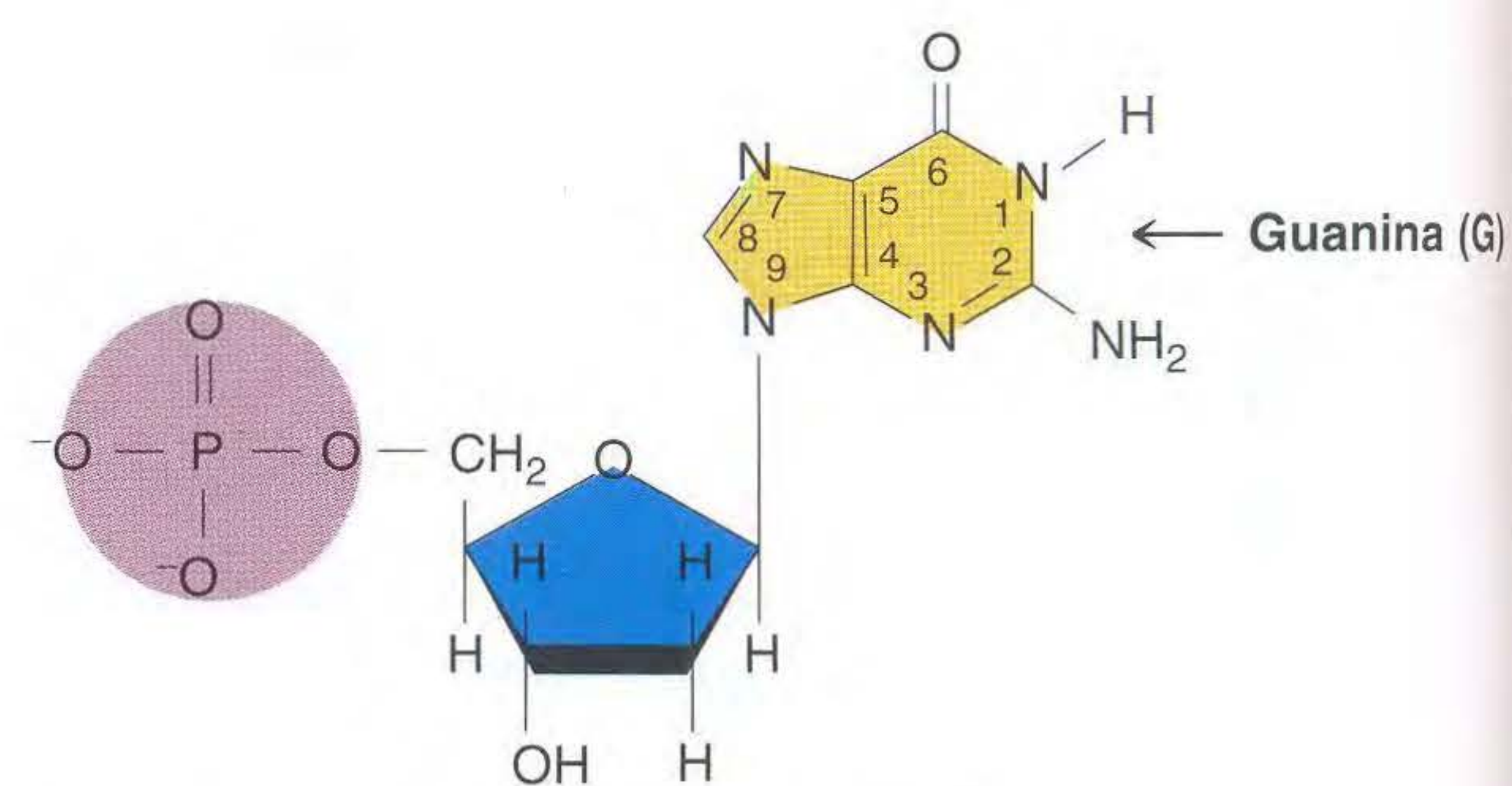
1. A quantidade total de nucleotídeos pirimidínicos (T + C) é sempre igual à quantidade total de nucleotídeos purínicos (A + G).
2. A quantidade de T é sempre igual à quantidade de A, e a quantidade de C é sempre igual à quantidade de G. Mas a quantidade de A + T não é necessariamente igual à quantidade de G + C, como pode ser visto na coluna direita do Quadro 7.1. Essa proporção varia entre organismos diferentes, mas é praticamente a mesma em tecidos diferentes do mesmo organismo.

Estruturas dos quatro nucleotídeos do DNA

Nucleotídeos purínicos

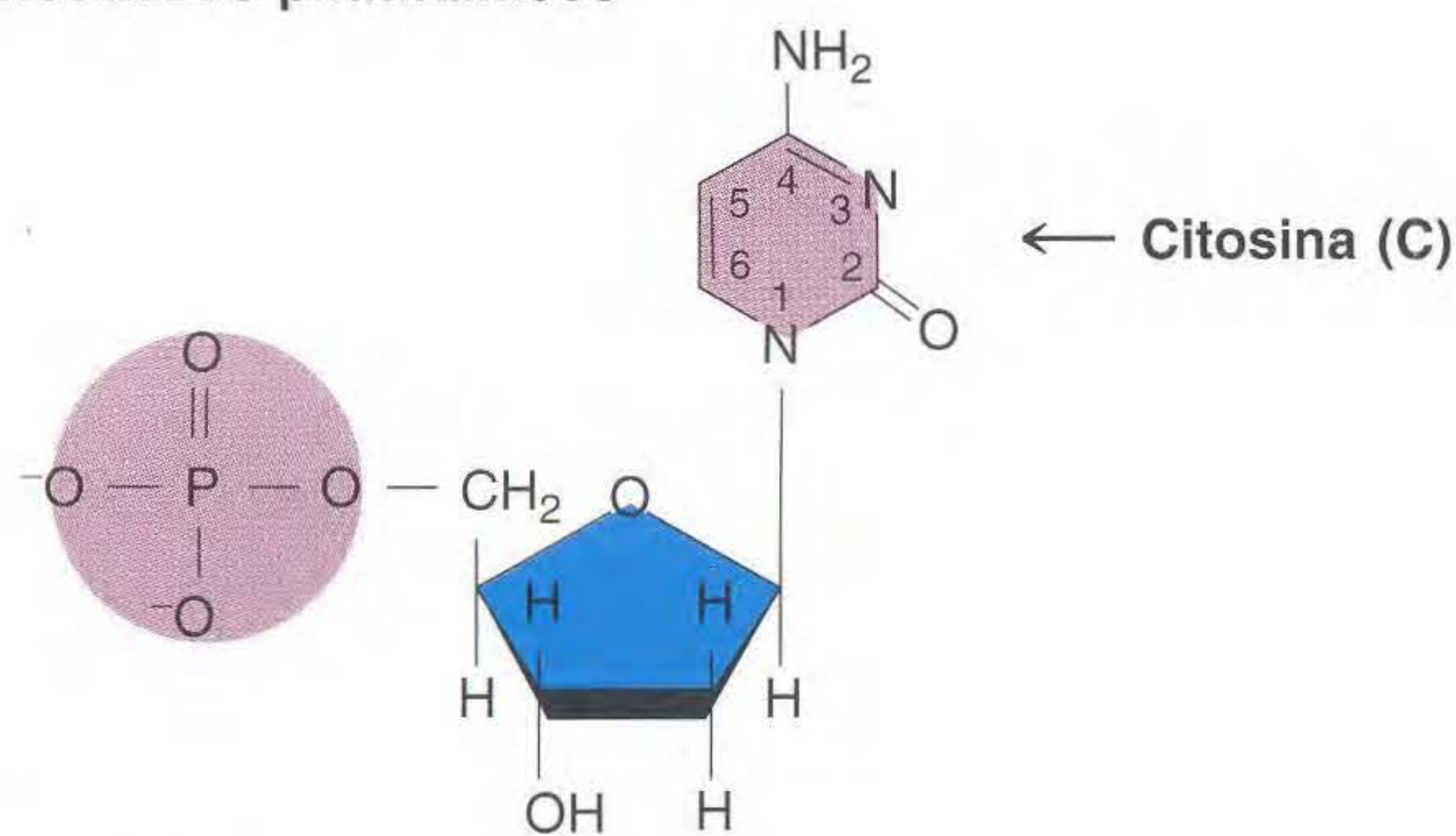


Desoxiadenosina 5'-monofosfato (dAMP)

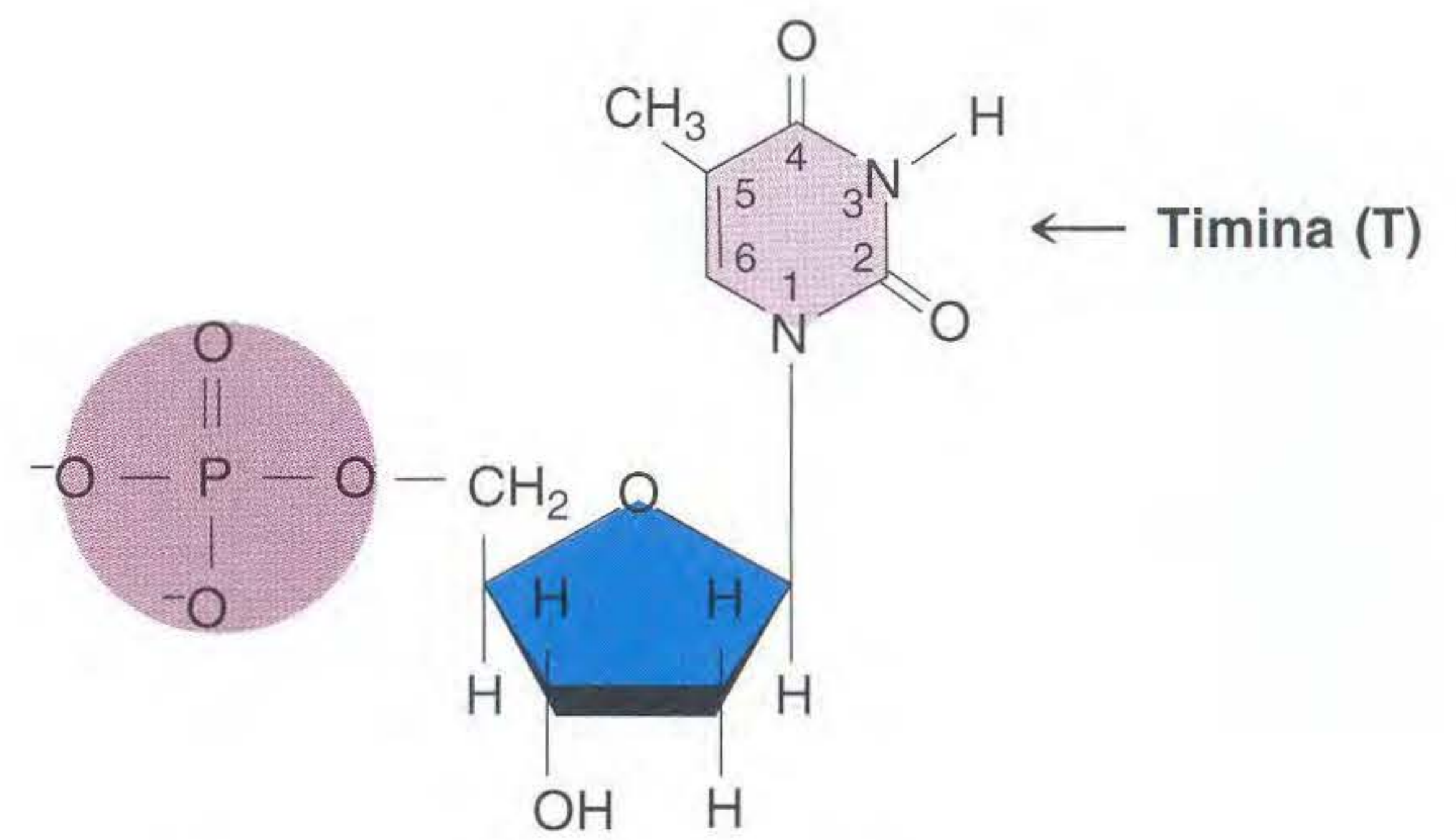


Desoxiguanosina 5'-monofosfato (dGMP)

Nucleotídeos pirimidínicos



Desoxicitidina 5'-monofosfato (dCMP)



Desoxitimidina 5'-monofosfato (dTMP)

FIG. 7.5 Esses nucleotídeos, dois com bases purínicas e dois com bases pirimidínicas, são os blocos fundamentais do DNA. O açúcar é chamado de *desoxirribose* porque é uma variação de um açúcar comum, *ribose*, que tem um átomo de oxigênio a mais (posição indicada pela seta vermelha).

Quadro 7.1 Propriedades Molares das Bases* em DNA de Várias Fontes

Organismo	Tecido	Adenina	Timina	Guanina	Citosina	$\frac{A+T}{G+C}$
<i>Escherichia coli</i> (K12)	-	26,0	23,9	24,9	25,2	1,00
<i>Diplococcus pneumoniae</i>	-	29,8	31,6	20,5	18,0	1,59
<i>Mycobacterium tuberculosis</i>	-	15,1	14,6	34,9	35,4	0,42
Levedura	-	31,3	32,9	18,7	17,1	1,79
<i>Paracentrotus lividus</i> (ouriço-do-mar)	Espermatozóide	32,8	32,1	17,7	18,4	1,85
Arenque	Espermatozóide	27,8	27,5	22,2	22,6	1,23
Rato	Medula óssea	28,6	28,4	21,4	21,5	1,33
Humano	Timo	30,9	29,4	19,9	19,8	1,52
Humano	Fígado	30,3	30,3	19,5	19,9	1,53
Humano	Espermatozóide	30,7	31,2	19,3	18,8	1,62

*Definidas em moles de constituintes nitrogenados por 100 g de átomos de fosfato em hidrolisados.

Fonte: E. Chargaff and J. Davidson, eds., *The Nucleic Acids*. Academic Press, 1955.

Resultado do experimento crítico de Rosalind Franklin

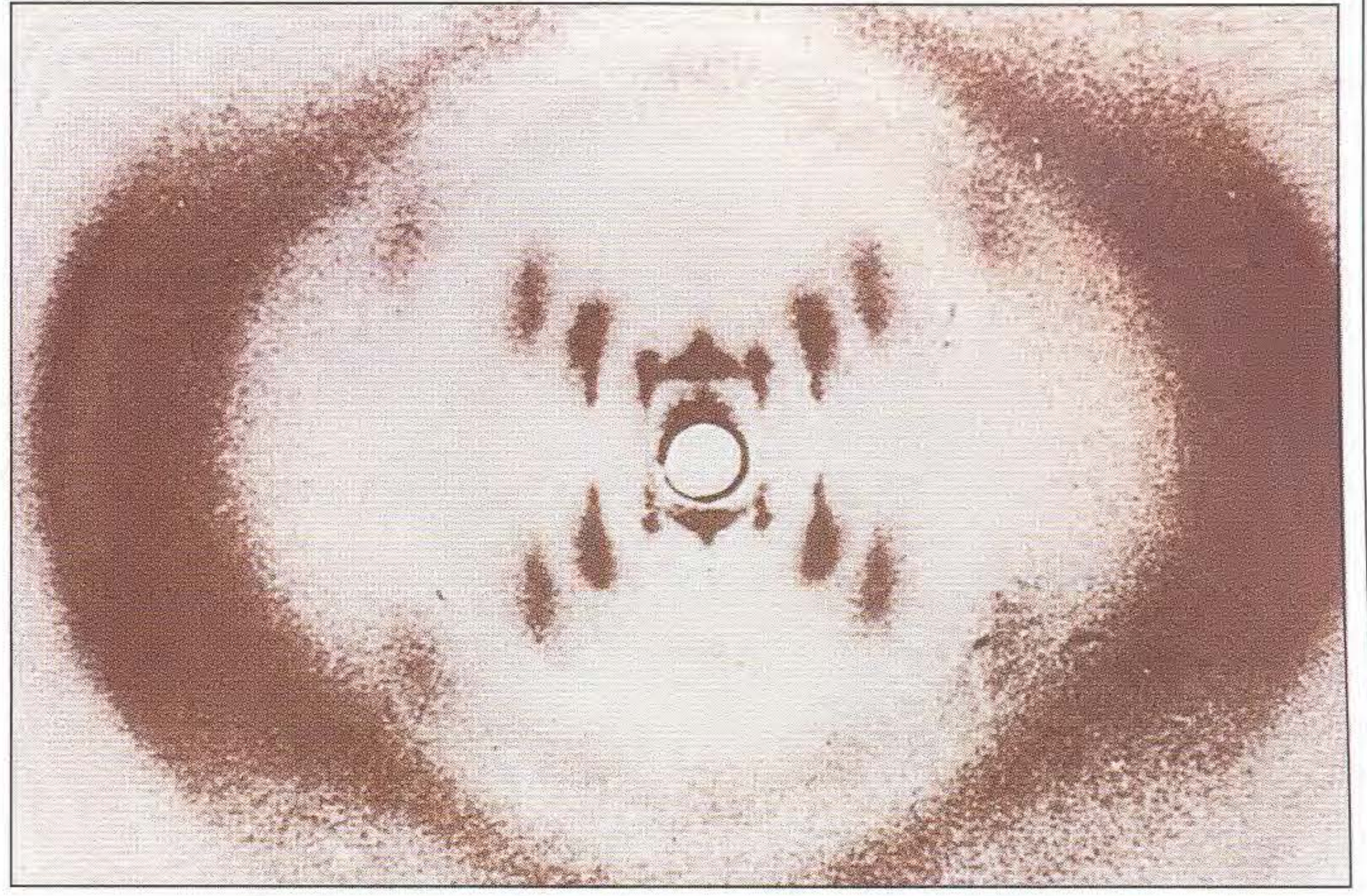


FIG. 7.6 Rosalind Franklin (esquerda) e seu padrão de difração de raios X do DNA (direita). [(Esquerda) Cortesia da National Portrait Gallery, London; (direita) Rosalind Franklin/Science Source/Photo Researchers.]

Análise de difração de raios X do DNA A terceira e mais controversa peça do quebra-cabeças veio dos dados de difração de raios X na estrutura do DNA que foram coletados por Rosalind Franklin quando ela estava no laboratório de Maurice Wilkins (Figura 7.6). Em tais experimentos, os raios X são disparados nas fibras de DNA, e a dispersão dos raios nas fibras é observada captando os raios em um filme fotográfico, no qual os raios X produzem pontos. O ângulo de dispersão representado por cada ponto no filme dá informações sobre a posição de um átomo ou alguns grupos de átomos na molécula de DNA. Esse procedimento não é simples de fazer (ou de explicar), e a interpretação dos padrões de pontos requer um tratamento matemático complexo que está além do escopo deste texto. Os dados disponíveis sugerem que o DNA é longo e fino, e tem duas partes similares que são paralelas umas às outras e correm ao longo da molécula. Os dados de raios X mostraram que a molécula é helicoidal (em espiral). Sem que Rosalind Franklin soubesse, sua melhor imagem de raios X foi mostrada a Watson e Crick por Maurice Wilkins, e essa era a peça principal do quebra-cabeças que lhes permitiu deduzir a estrutura tridimensional que podia explicar os padrões de pontos nos raios X.

A dupla hélice

Uma publicação de 1953 de Watson e Crick no periódico *Nature* começa com duas frases que abriram uma nova era da Biologia: “Nós queremos sugerir uma estrutura para o sal de ácido desoxirribonucleico (D.N.A.). Essa estrutura tem novas características que são de considerável interesse biológico.” A estrutura do DNA foi objeto de um grande debate desde os experimentos de Avery *et al.* em 1944. Como vimos, a composição geral do DNA era conhecida, mas como suas partes se reuniam não era conhecido. A estrutura preenchia os principais requisitos de uma molécula hereditária: a habilidade em estocar informação, a habilidade em se replicar e a habilidade em mutar.

A estrutura tridimensional decifrada por Watson e Crick é composta de duas cadeias lado a lado (“filamentos”) de nucleotídeos torcidos na forma de uma **dupla hélice** (Figura 7.7). Os dois filamentos de nucleotídeos são mantidos juntos por pontes de hidrogênio entre as bases de cada filamento, formando uma estrutura como uma escada em espiral (Figura 7.8a). O arcabouço de cada filamento é formado de unidades alternadas de fosfato e desoxirribose que são conectadas por ligações fosfodiéster (Figura 7.8b). Podemos usar essas ligações para descrever como uma cadeia de nucleotídeos é organizada. Como já mencionado, os átomos de carbono dos grupamentos açúcar são numerados de 1' a 5'. Uma ligação fosfodiéster conecta o átomo de carbono 5' de uma desoxirribose ao átomo de carbono 3' da desoxirribose adjacente. Assim, cada ligação açúcar-fosfato é dita como tendo uma polaridade 5' para 3', ou sentido, e a compreensão dessa polaridade é essencial na compreensão de como o DNA pre-

O primeiro modelo de DNA



FIG. 7.7 James Watson e Francis Crick com seu modelo do DNA. [Camera Press.]

A estrutura do DNA

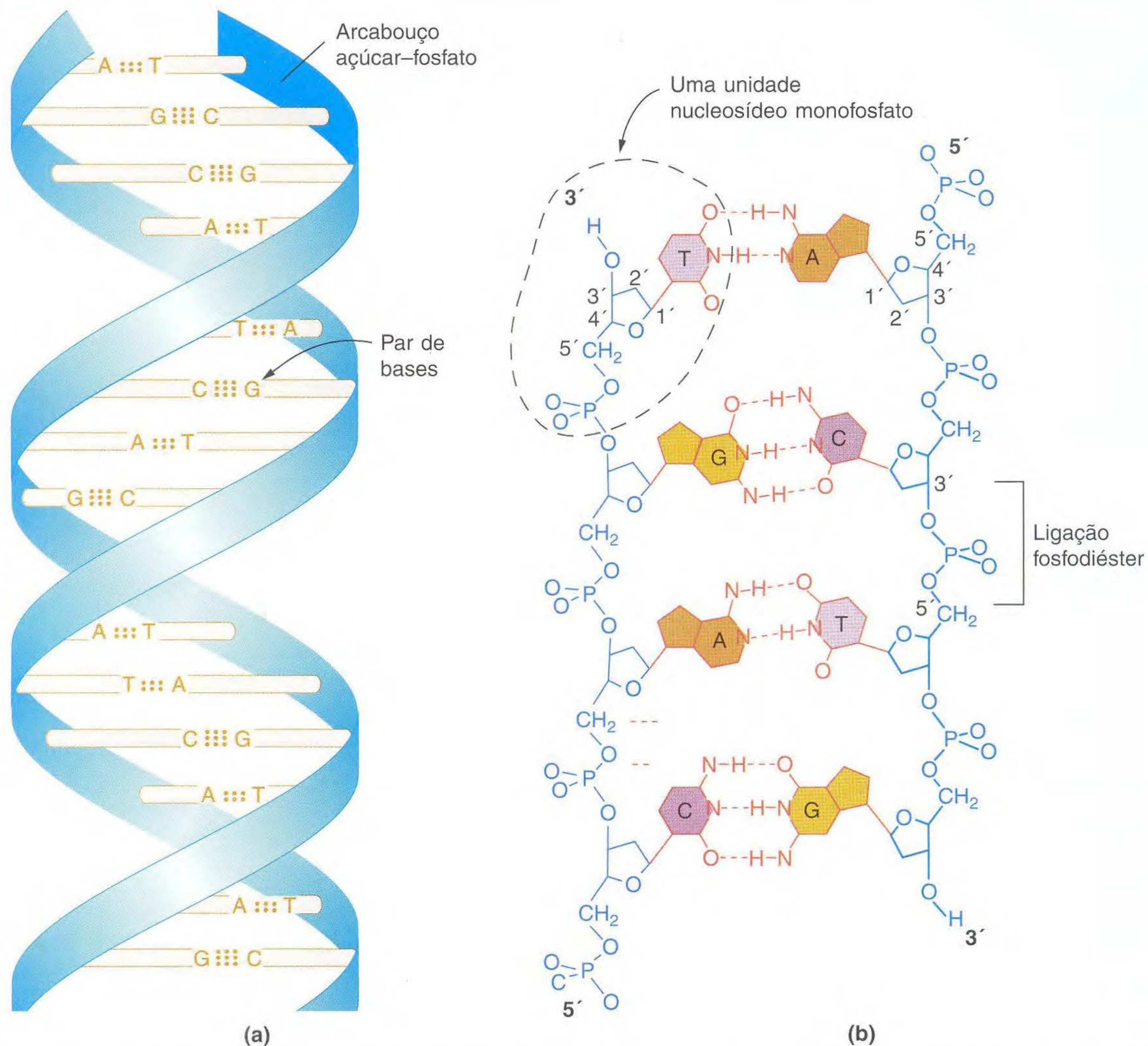


FIG. 7.8 (a) Modelo simplificado mostrando a estrutura helicoidal do DNA. Os bastões representam pares de bases, e as fitas representam os arcabouços açúcar-fosfato das duas cadeias de polaridade inversa. (b) Um diagrama químico preciso da dupla hélice de DNA, desenrolado para mostrar os arcabouços açúcar-fosfato (azul) e degraus de pares de bases (rosa). Os arcabouços correm em sentidos opostos; as pontas 5' e 3' são denominadas pela orientação dos átomos de carbono 5' e 3' dos anéis de açúcar. Cada par de bases tem uma base purina, adenina (A) ou guanina (G), e uma base pirimidínica, timina (T) ou citosina (C), conectadas por pontes de hidrogênio (linhas tracejadas). [De R.E. Dickerson, "The DNA Helix and How It is Read." Copyright 1983 by Scientific American, Inc. Todos os direitos reservados.]

enche seus papéis. Na molécula bifilamentar de DNA, os dois arcabouços estão em orientação oposta, ou **em polaridade inversa** (veja Figura 7.8b).

Cada base é ligada ao átomo de carbono 1' de um açúcar desoxirribose no arcabouço de cada filamento e está voltada para dentro até uma base no outro filamento. As pontes de hidrogênio entre os pares de bases devem manter os dois filamentos da molécula de DNA juntos. As pontes de hidrogênio são indicadas por linhas tracejadas na Figura 7.8b.

Os dois filamentos nucleotídicos pareados com polaridade inversa automaticamente assumem uma conformação de dupla hélice (Figura 7.9), principalmente pela interação de pares de bases. Os pares de bases, que são estruturas planares achatadas, empilham-se uns sobre os outros no centro da

dupla hélice (veja Figura 7.9a). O empilhamento dá estabilidade à molécula de DNA, excluindo as moléculas de água dos espaços entre os pares de bases. A forma mais estável que resulta do empilhamento de bases é uma dupla hélice com dois tamanhos distintos de sulcos correndo em uma hélice: o **sulco maior** e o **sulco menor**, como mostram ambas as fitas do modelo compactado (Figura 7.9b). A maioria das associações DNA-proteína são nos sulcos maiores. Um único filamento de nucleotídeos não tem estrutura helicoidal; a forma helicoidal do DNA depende totalmente do pareamento e empilhamento das bases nos filamentos de polaridade inversa. O DNA é uma hélice com giro para a direita; em outras palavras, ele tem a mesma estrutura de um parafuso que seria aparafusado no lugar usando um movimento de giro horário.

Duas representações da dupla hélice de DNA

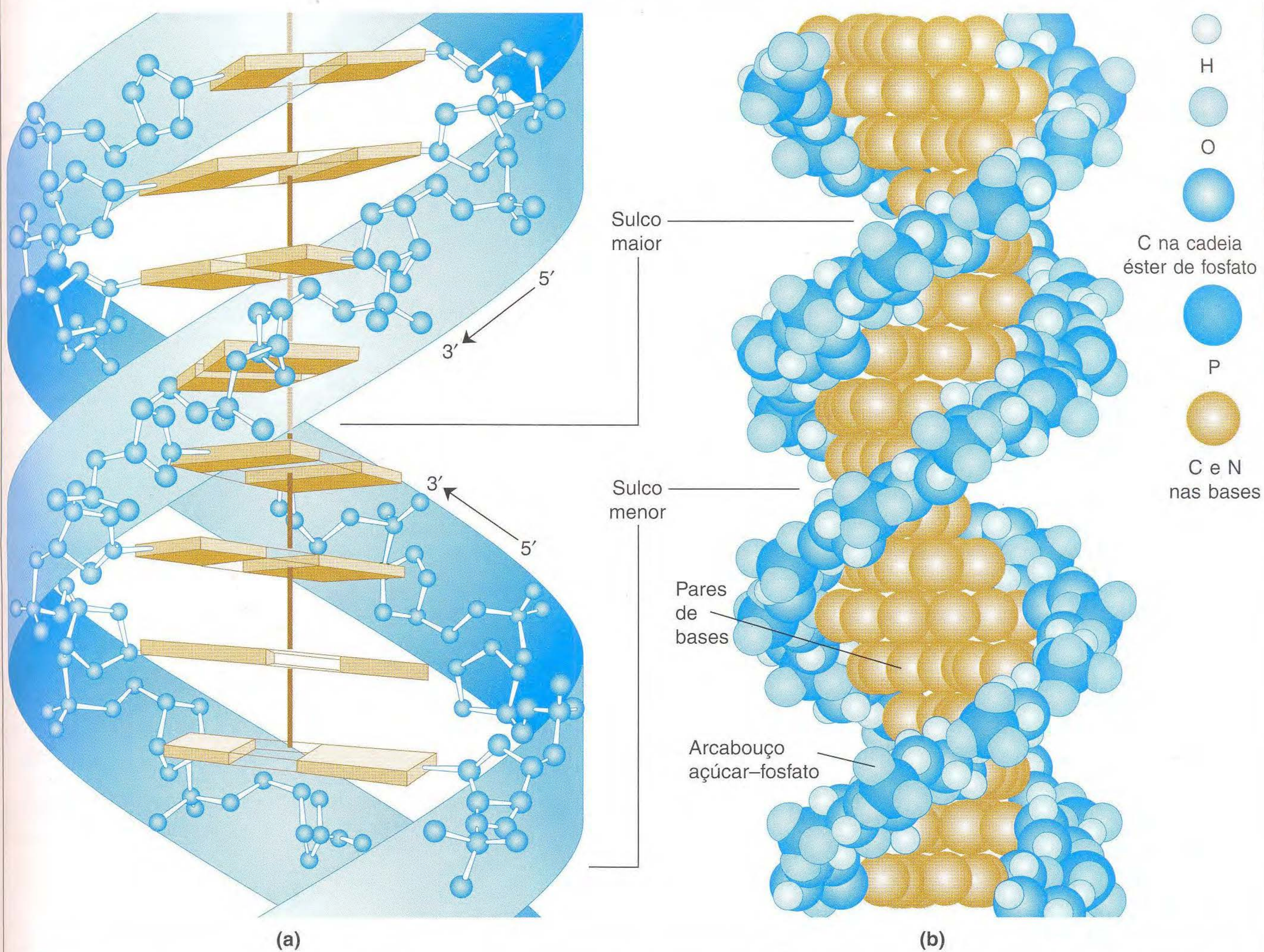


FIG. 7.9 O diagrama em fita (a) destaca o empilhamento de pares de bases, enquanto o modelo compactado (b) mostra os sulcos maior e menor. [(b) De C. Yanofsky, "Gene Structure and Protein Structure." Copyright 1967 by Scientific American, Inc. Todos os direitos reservados.]

A dupla hélice responde muito bem aos dados de raios X e, bem-sucedidamente, aos dados de Chargaff. Estudando os modelos que eles fizeram da estrutura, Watson e Crick perceberam que o raio observado da dupla hélice (conhecido dos dados de raios X) seriam explicados se uma purina sempre fizesse par (por pontes de hidrogênio) com uma base pirimidina (Figura 7.10). Tal pareamento responderia pela regularidade $(A + G) = (T + C)$ observada por Chargaff, mas seriam previstos quatro possíveis pareamentos: $T \cdots A$, $T \cdots G$, $C \cdots A$ e $C \cdots G$. Os dados de Chargaff, entretanto, indicam que T só pareia com A, e C só pareia com G. Watson e Crick concluíram que cada par de bases consistia em uma base purina e uma base pirimidina, pareadas de acordo com a seguinte regra: G pareia com C e A pareia com T.

Note que o par G-C tem três pontes de hidrogênio, enquanto o par A-T só tem duas (veja Figura 7.8b). Seria previsto que o DNA contendo muitos pares G-C seria mais estável que o DNA contendo muitos pares A-T. De fato, essa previsão é confirmada. O calor faz com que os dois filamentos de DNA

Pareamento de bases no DNA

Pirimidina + pirimidina:
DNA muito fino

Purina + purina:
DNA muito grosso

Purina + pirimidina: espessura
compatível com dados de raios X

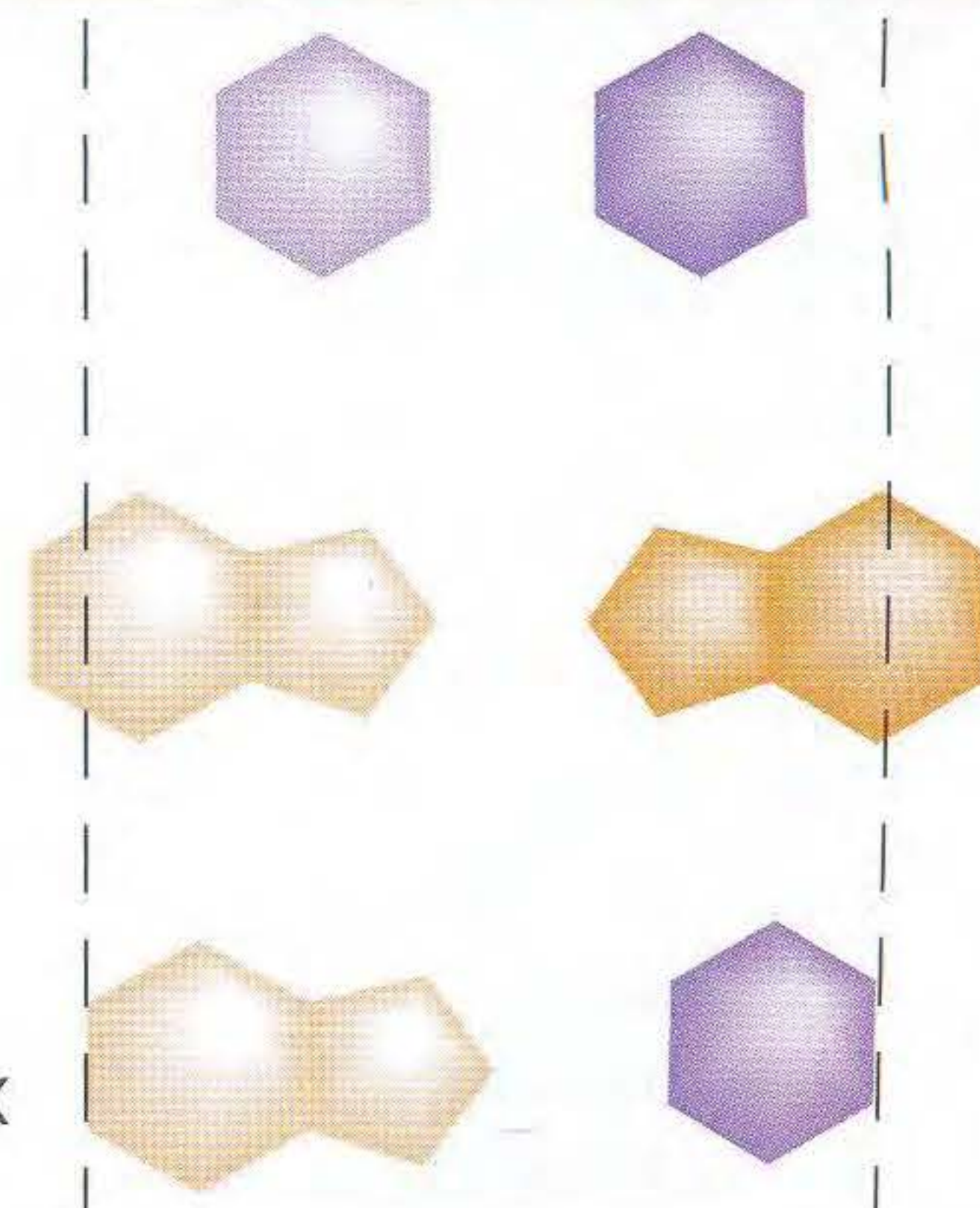


FIG. 7.10 O pareamento de purinas com pirimidinas responde exatamente pelo diâmetro da dupla hélice de DNA determinada por dados de raios X. Esse diâmetro é indicado pelas linhas verticais tracejadas. [De R.E. Dickerson, "The DNA Helix and How It Is Read." Copyright 1983 by Scientific American, Inc. Todos os direitos reservados.]

se separem (um processo chamado de dissociação (*melting*) do DNA ou desnaturação do DNA). Os DNA com conteúdo maior de G + C podem ser mostrados requerendo maiores temperaturas para se dissociar devido à maior atração do pareamento G-C.

O DNA é uma dupla hélice composta de duas cadeias de nucleotídeos mantidas juntas pelo pareamento complementar de A com T e G com C.

A descoberta de Watson e Crick da estrutura do DNA é considerada por alguns como sendo a descoberta biológica mais importante do século XX, e mereceu o Prêmio Nobel para Maurice Wilkins em 1962 (Rosalind Franklin morreu de câncer em 1958, e o prêmio não é concedido após a morte). O motivo de essa descoberta ser considerada tão importante é que a dupla hélice, além de ser consistente com os dados anteriores sobre a estrutura do DNA, preenche os três requisitos para uma substância hereditária.

1. A estrutura em dupla hélice sugere como o material genético pode determinar a estrutura das proteínas. Talvez a *seqüência* de pares de nucleotídeos no DNA determine a seqüência de aminoácidos na proteína especificada por esse gene. Em outras palavras, algum tipo de **código genético** pode escrever uma informação no DNA como uma seqüência de nucleotídeos e, então, traduzi-la em uma linguagem diferente de seqüências de aminoácidos na proteína. Como isso é feito é o assunto do Capítulo 9.
2. Se a seqüência de bases do DNA especifica a seqüência de aminoácidos, então é possível uma mutação pela substituição de um tipo de base por outro em uma ou mais posições. As mutações serão discutidas no Capítulo 15.
3. Como Watson e Crick disseram nas palavras de conclusão de sua publicação de 1953 na *Nature*, relatando a estrutura de dupla hélice do DNA: "Não escapou à nossa percepção que o pareamento específico que postulamos sugere imediatamente um possível mecanismo de cópia para o material genético." Para os geneticistas da época, o significado dessa afirmativa estava claro, como veremos na seção seguinte.

► 7.3 Replicação Semiconservativa

O mecanismo de cópia ao qual Watson e Crick se referiram é chamado de replicação semiconservativa, e está diagramado na Figura 7.11. Os arcabouços açúcar-fosfato são representados por fitas espessas, e a seqüência de pares de bases é aleatória. Imaginemos que a dupla hélice é análoga a um zíper que se abre, começando em uma ponta (veja Figura 7.11). Podemos ver que, se essa analogia com o zíper é válida, a deslicoidização dos dois filamentos irá expor as bases em cada filamento. Cada base exposta tem o potencial para parear com nucleotídeos livres na solução. Como a estrutura do DNA impõe requisitos rígidos de pareamento, cada base exposta irá parear apenas com sua **base complementar**, A com T e G com C. Assim, cada um dos dois filamentos irá agir como um **molde**, para dirigir a montagem de bases comple-

Replicação semiconservativa do DNA

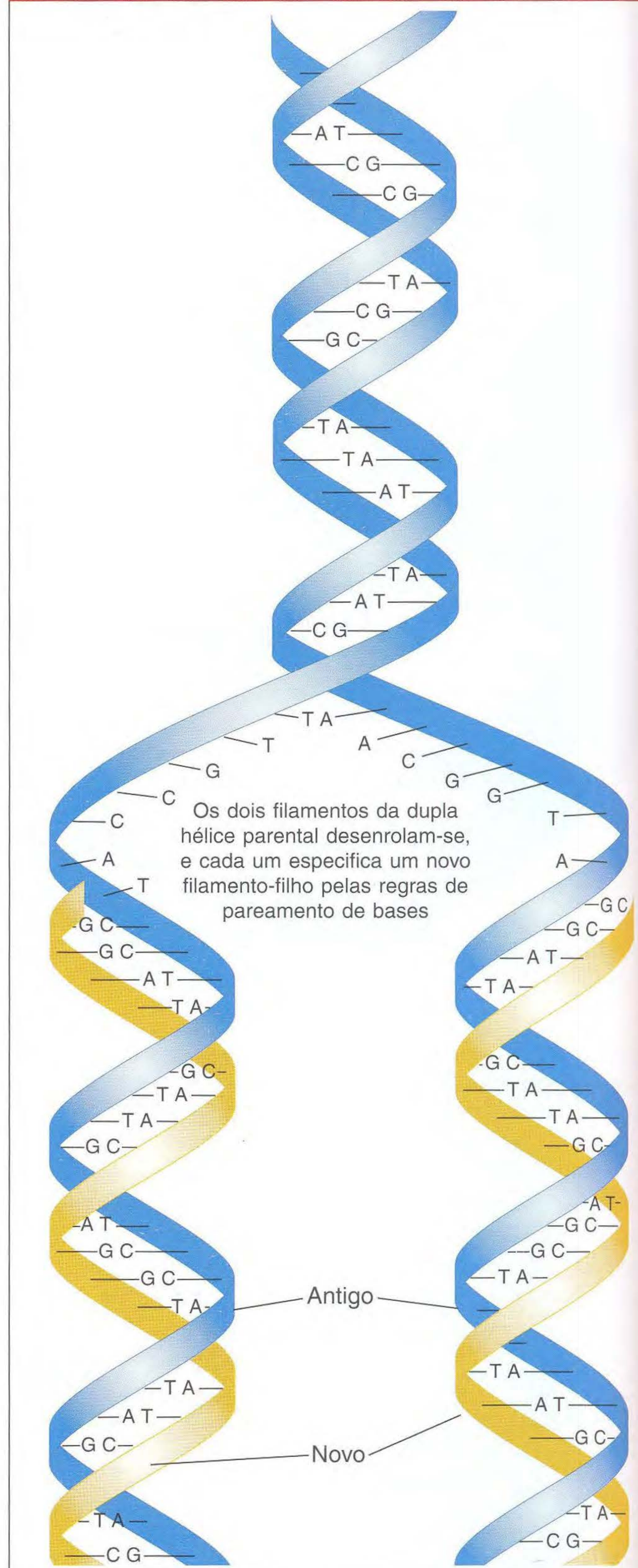


FIG. 7.11 O modelo semiconservativo da replicação do DNA proposto por Watson e Crick é baseado na especificidade de pontes de hidrogênio dos pares de bases. Os filamentos parentais, mostrados em azul, servem como moldes para a polimerização. Os filamentos recém-polimerizados, mostrados em amarelo, têm seqüências de bases que são complementares a seus respectivos moldes.

mentares para reestruturar uma dupla hélice idêntica à original. Os nucleotídeos recém-adicionados são supostos como vindos de um conjunto de nucleotídeos livres que devem estar presentes na célula.

Se esse modelo está correto, então cada molécula-filha deve conter uma cadeia parental de nucleotídeos e uma recém-sintetizada. Entretanto, pensando um pouco, vemos que existem pelo menos três modos diferentes pelos quais a molécula de DNA parental pode estar relacionada às moléculas-filhas. Esses modos hipotéticos de replicação são chamados de semiconservativo (o modelo de Watson e Crick), conservativo e dispersivo (Figura 7.12). Na **replicação semiconservativa**, a dupla hélice de cada molécula-filha de DNA contém um filamento da molécula original de DNA e um filamento recém-sintetizado. Entretanto, na **replicação conservativa**, a molécula parental de DNA é conservada, e uma única dupla hélice-filha é produzida, consistindo em dois filamentos recém-sintetizados. Na **replicação dispersiva**, as moléculas-filhas consistem em filamentos, cada um contendo segmentos de *ambos*, DNA parentais e DNA recém-sintetizados.

Experimento de Meselson-Stahl

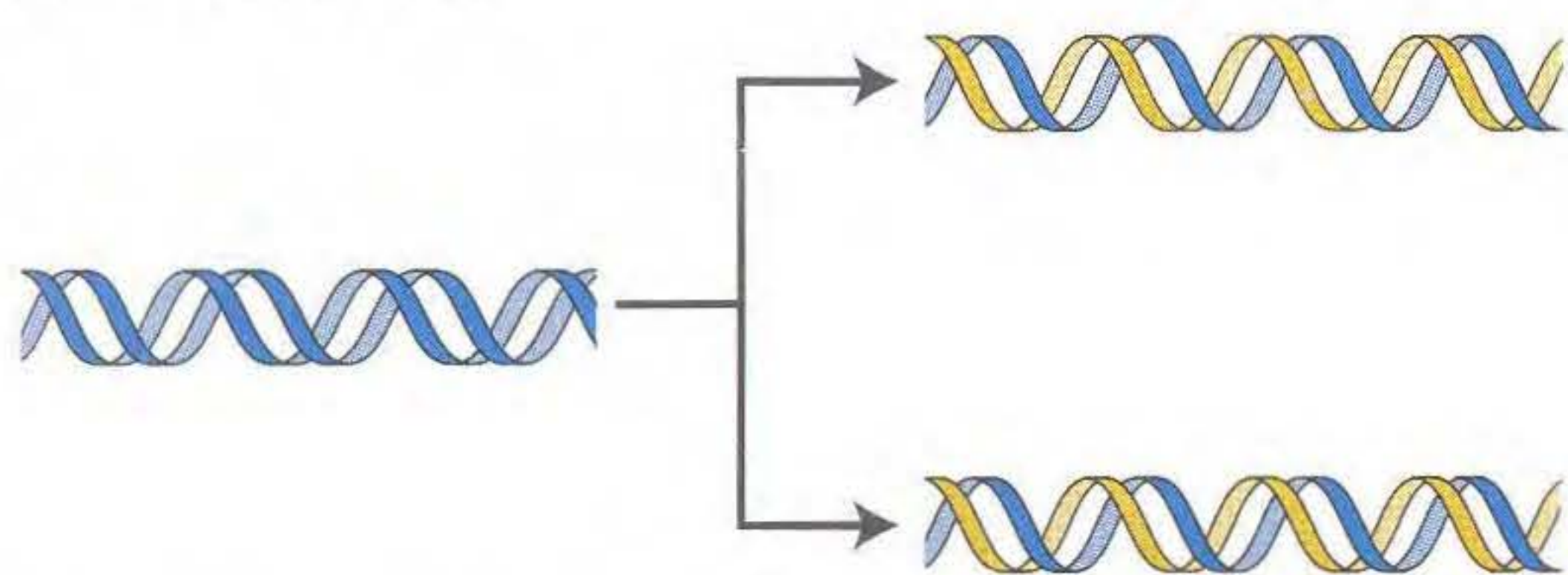
O primeiro problema na compreensão da replicação do DNA foi saber se o mecanismo de replicação era semiconservativo, conservativo ou dispersivo. Em 1958, dois jovens cientistas, Matthew Meselson e Franklin Stahl, descobriram qual

dessas possibilidades corretamente descreve a replicação do DNA. A idéia deles era permitir que as moléculas de DNA parental contendo nucleotídeos de uma densidade fossem replicadas em um meio contendo nucleotídeos de diferentes densidades. Se o DNA fosse replicado semiconservativamente, as moléculas-filhas deveriam ser metade velhas e metade novas, e, portanto, de densidade intermediária. Para fazer seu experimento, eles cultivaram células de *E. coli* em um meio contendo o isótopo pesado de nitrogênio (^{15}N) ao invés da forma leve normal (^{14}N). Esse isótopo foi inserido nas bases nitrogenadas, que foram então incorporadas em filamentos de DNA recém-sintetizados. Após muitas divisões celulares em ^{15}N , o DNA das células foi bem marcado com o isótopo pesado. As células foram então removidas do meio com ^{15}N e colocadas em um meio com ^{14}N . Após uma ou duas divisões celulares, as amostras foram colhidas e o DNA foi isolado de cada amostra.

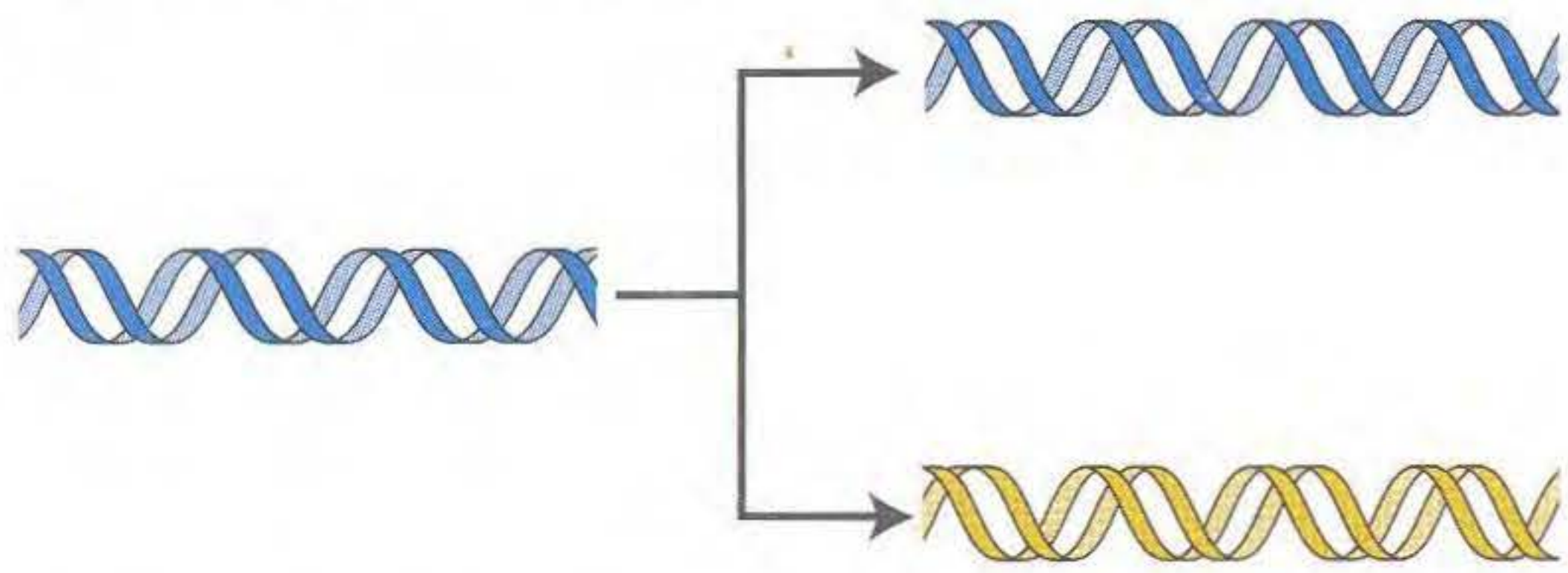
Meselson e Stahl foram capazes de distinguir o DNA de densidades diferentes porque as moléculas podem ser separadas umas das outras por um procedimento chamado *centrifugação em gradiente de cloreto de cério*. Se o cloreto de cério (CsCl) for centrifugado em velocidades altíssimas (50.000 rpm) por muitas horas, os íons de cério e cloreto tendem a ser levados pela força centrífuga para o fundo do tubo. Finalmente, é estabelecido um gradiente de íons no tubo, com a concentração mais alta de íons, ou densidade, no fundo. O DNA centrifugado com cloreto de cério forma uma banda em uma posição idêntica com sua densidade no gradiente (Figura 7.13). O DNA de densidades diferentes formará bandas em locais diferentes. As células inicialmente cultivadas no isótopo pesado ^{15}N mostraram DNA de alta densidade. Esse DNA é mostrado em azul no lado esquerdo da Figura 7.13a. Após cultivar essas células no isótopo leve ^{14}N por uma geração, os pesquisadores encontraram que o DNA era de densidade intermediária, mostrado metade em azul (^{15}N) e metade em amarelo (^{14}N) na metade da Figura 7.13a. Note que Meselson e Stahl continuaram o experimento por duas gerações de *E. coli*, de modo que eles podiam distinguir a replicação semiconservativa da dispersiva. Após duas gerações, ambos os DNA, intermediário e de baixa densidade, foram observados (lado direito da Figura 7.13a), exatamente como o previsto pelo modelo semiconservativo de replicação de Watson-Crick.

Três modelos alternativos para replicação do DNA

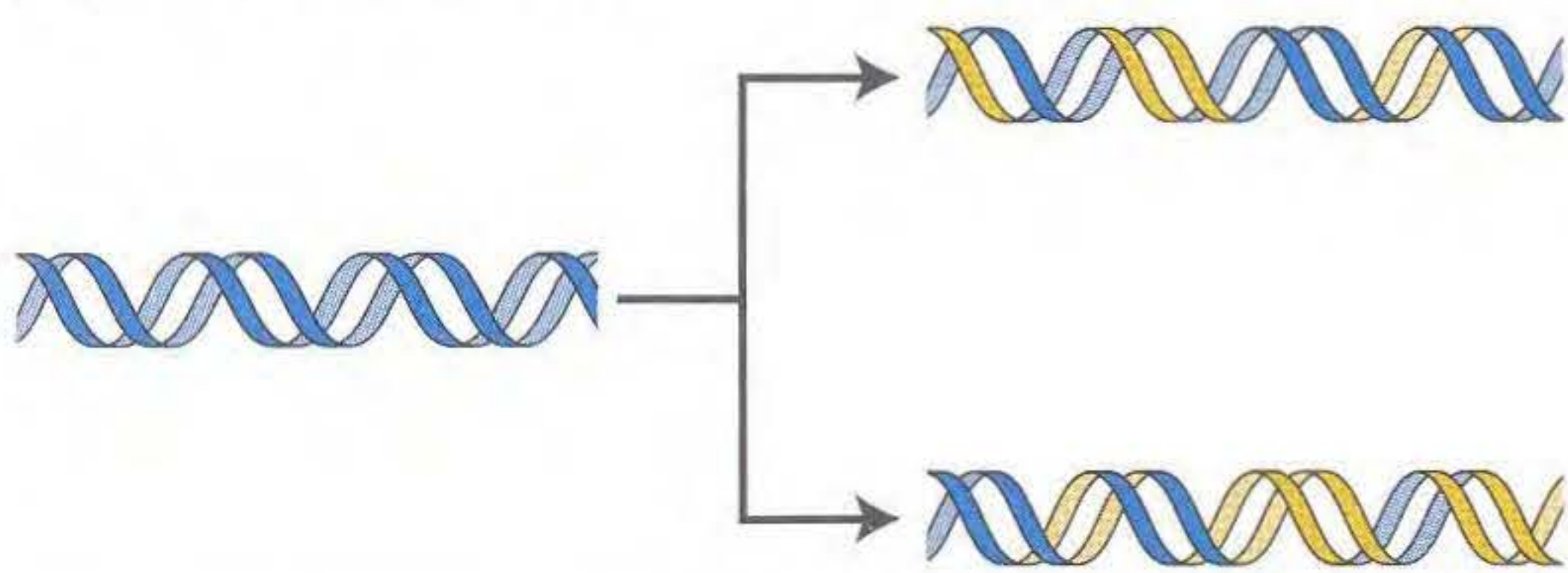
Replicação semiconservativa



Replicação conservativa



Replicação dispersiva



O DNA é replicado pela deslicoidização dos dois filamentos da dupla hélice e construção de filamentos complementares a cada um dos filamentos separados da dupla hélice original.

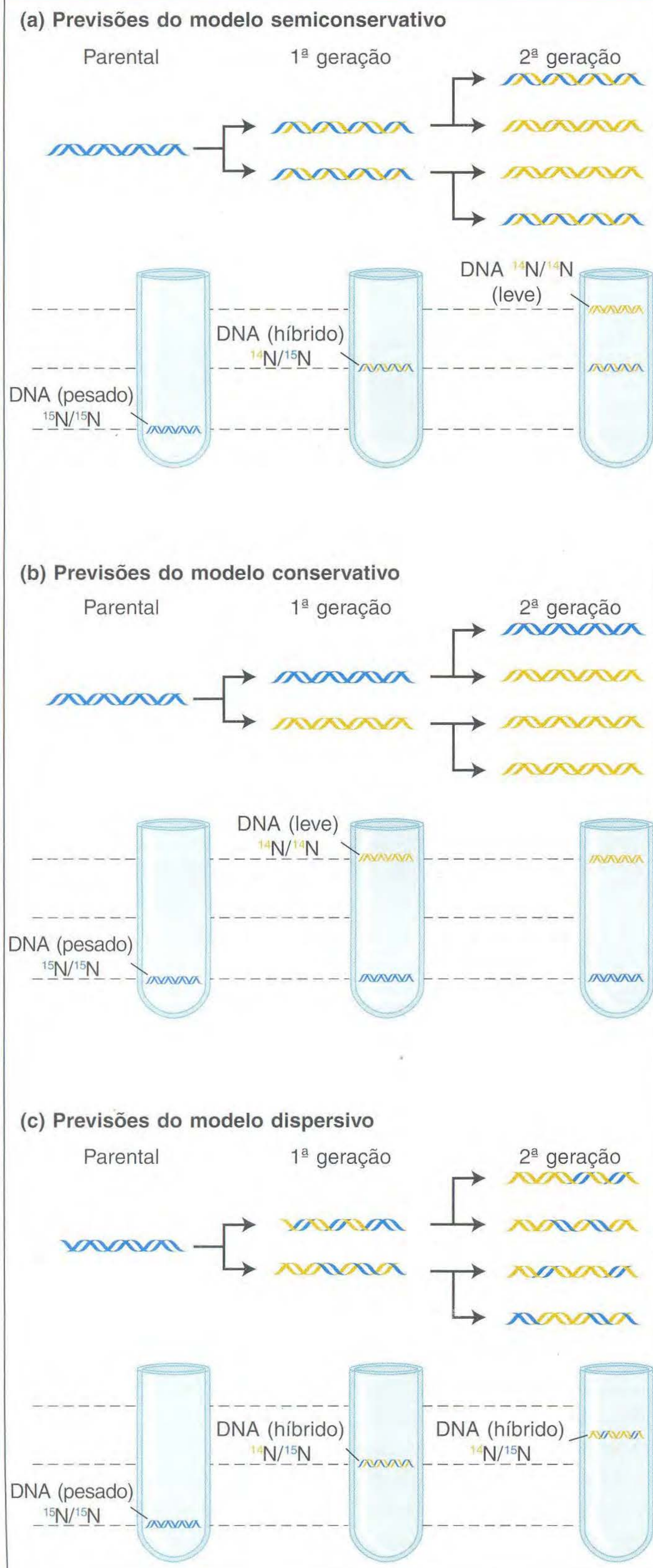
A forquilha de replicação

Outra previsão do modelo de Watson-Crick da replicação do DNA é que o zíper de replicação, ou *forquilha*, seria encontrado na molécula de DNA durante a replicação. Essa forquilha é o local no qual a dupla hélice é desenrolada para produzir os dois filamentos únicos que servem como moldes para a cópia. Em 1963, John Cairns testou essa previsão permitindo a replicação do DNA nas células bacterianas para incorpo-

FIG. 7.12 Dos três modelos alternativos de replicação do DNA, o modelo de Watson-Crick da estrutura do DNA produziria o primeiro modelo (semiconservativo). As linhas amarelas representam os filamentos recém-sintetizados.

rar a timidina triciada ($[^3\text{H}]$ timidina), o nucleotídeo timina marcado com um isótopo radioativo de hidrogênio chamado trício. Teoricamente, cada nova molécula-filha sintetizada

O DNA é copiado pela replicação semiconservativa



deveria conter um filamento radioativo (“quente”) (com ^3H) e outro não-radioativo (“frio”). Após intervalos variados e números variados de ciclos de replicação em um meio “quente”, Cairns cuidadosamente lisou a bactéria e permitiu que os conteúdos da célula ficassem em um pedaço de papel de filtro, que foi posto em uma lâmina de microscópio. Finalmente, Cairns cobriu o filtro com uma emulsão fotográfica e a expôs no escuro por 2 meses. Esse procedimento, chamado auto-radiografia, permitiu que Cairns desenvolvesse uma imagem da localização de ^3H no material celular. À medida que há o decaimento de ^3H , ele emite partículas beta (um elétron energético). A emulsão fotográfica detecta uma reação química que ocorre sempre que uma partícula beta atinge a emulsão. A emulsão pode então ser revelada como uma fotografia, de modo que a trilha de emissão de partículas beta aparece como pontos pretos ou grãos.

Após um ciclo de replicação em $[^3\text{H}]$ timidina, aparece um anel de pontos na auto-radiografia. Cairns interpretou esse anel como um filamento radioativo recém-formado em uma molécula circular de DNA, como mostrado na Figura 7.14a. É, portanto, aparente que o cromossomo bacteriano é circular, um fato que também emergiu de uma análise genética já descrita (veja Capítulo 5). No segundo ciclo de replicação, as forquilhas previstas pelo modelo foram de fato vistas. Além disso, a densidade dos grãos nos três segmentos era tal que pôde ser feita a interpretação mostrada na Figura 7.14b: a curva espessa de pontos passando pelo interior do círculo de DNA seria o filamento-filho recém-sintetizado, dessa vez consistindo em *dois* filamentos radioativos. Cairns viu todos os tamanhos desses padrões auto-radiográficos em forma de lua, correspondendo ao movimento progressivo das forquilhas de replicação, ao redor do anel. As estruturas do tipo mostrado na Figura 7.14b são chamadas de **estruturas teta (θ)**.

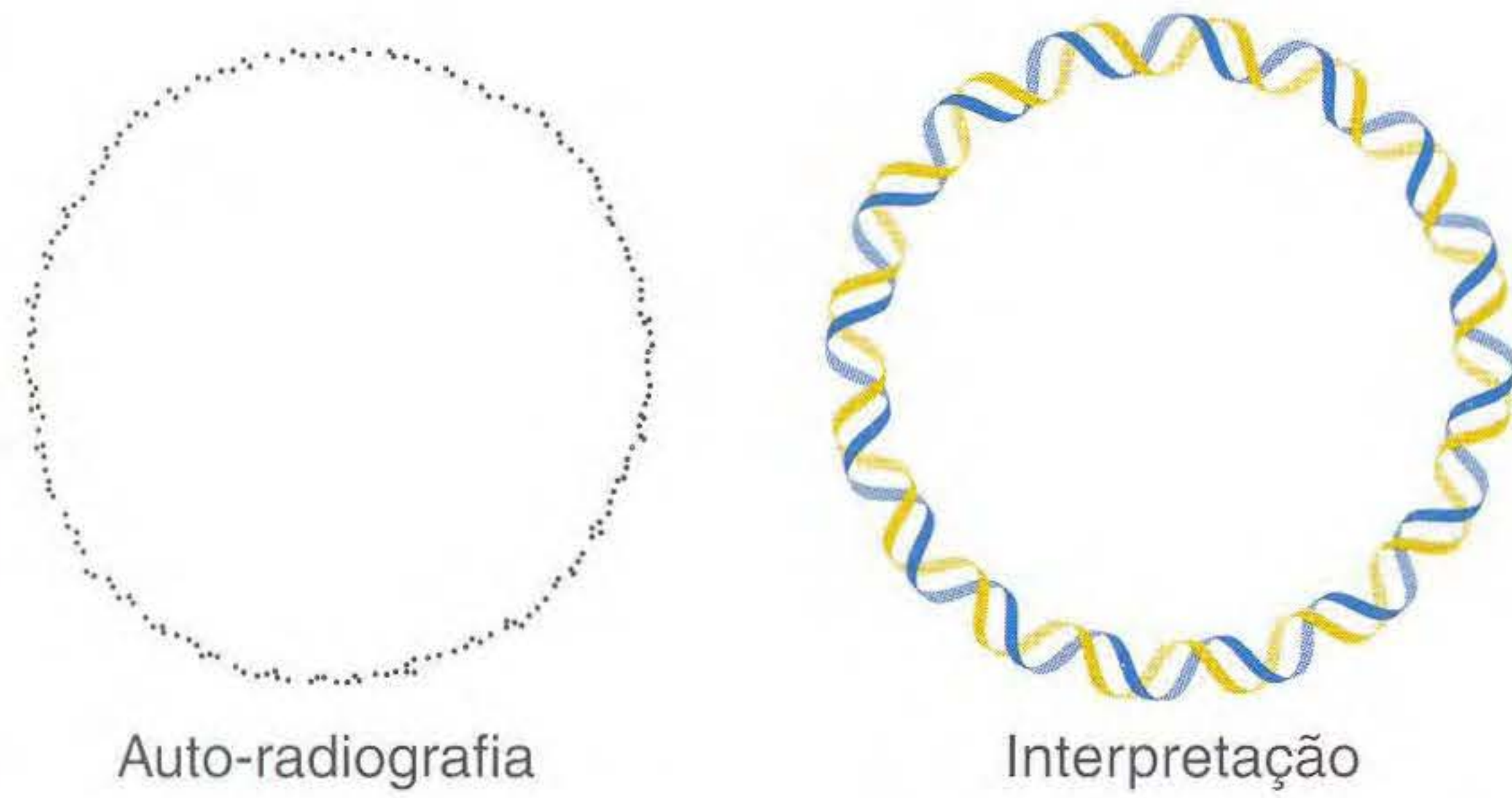
DNA polimerases

Um problema confrontado pelos cientistas foi compreender como essas bases são trazidas para a dupla hélice-molde. Embora os cientistas suspeitem que as enzimas tenham um papel, essa possibilidade não foi provada até 1959, quando Arthur Kornberg isolou a DNA polimerase de *E. coli* e demonstrou essa atividade enzimática *in vitro*. Essa enzima adiciona desoxirribonucleotídeos à ponta 3' de uma cadeia crescente de nucleotídeos, usando como molde um único filamento de DNA que foi exposto pela deselcoidização localizada da dupla hélice (Figura 7.15). Os substratos para a DNA polimerase são as formas trifosfato de desoxirribonucleotídeos, dATP, dGTP, dCTP e dTTP.

FIG. 7.13 O experimento de Meselson-Stahl demonstra que o DNA é copiado por replicação semiconservativa. O DNA centrifugado em gradiente de cloreto de cério (CsCl) formará bandas de acordo com sua densidade. (a) Quando as células cultivadas em ^{15}N são transferidas para um meio com ^{14}N , a primeira geração produz uma única banda de DNA intermediário, e a segunda geração produz duas bandas, uma intermediária e uma leve. Esse resultado é compatível com as previsões do modelo semiconservativo da replicação do DNA. (b e c) Os resultados previstos para a replicação conservativa e dispersiva, mostrados aqui, *não* foram encontrados.

Um cromossomo bacteriano replicante

(a) Cromossomo após uma rodada de replicação



(b) Cromossomo durante a segunda rodada de replicação

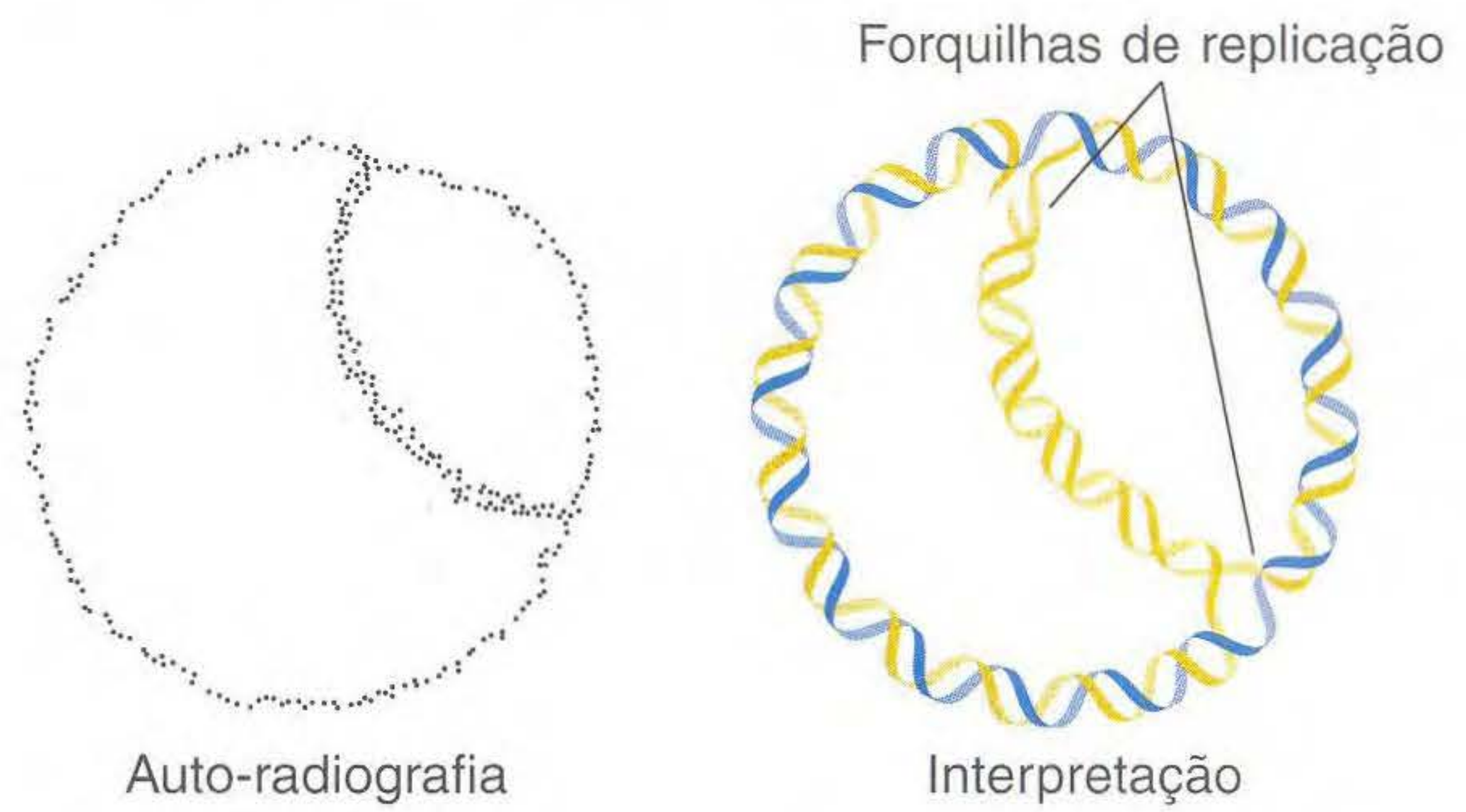


FIG. 7.14 Um cromossomo bacteriano replicante tem duas forquilha de replicação. (a) *Esquerda*: Auto-radiografia de um cromossomo bacteriano após uma replicação em timidina triciada. De acordo com o modelo semiconservativo da replicação, um dos dois filamentos deve ser radioativo. *Direita*: Interpretação da auto-radiografia. A hélice amarela representa o filamento triciado. (b) *Esquerda*: Auto-radiografia de um cromossomo bacteriano na segunda rodada de replicação em timidina triciada (^3H). Nessa estrutura teta (θ), a dupla hélice recém-replicada que cruza o círculo poderia consistir em dois filamentos radioativos (se o filamento parental fosse o radioativo). *Direita*: A espessura dupla do traçador radioativo no auto-radiograma parece confirmar a interpretação mostrada aqui.

Reação catalisada pela DNA polimerase

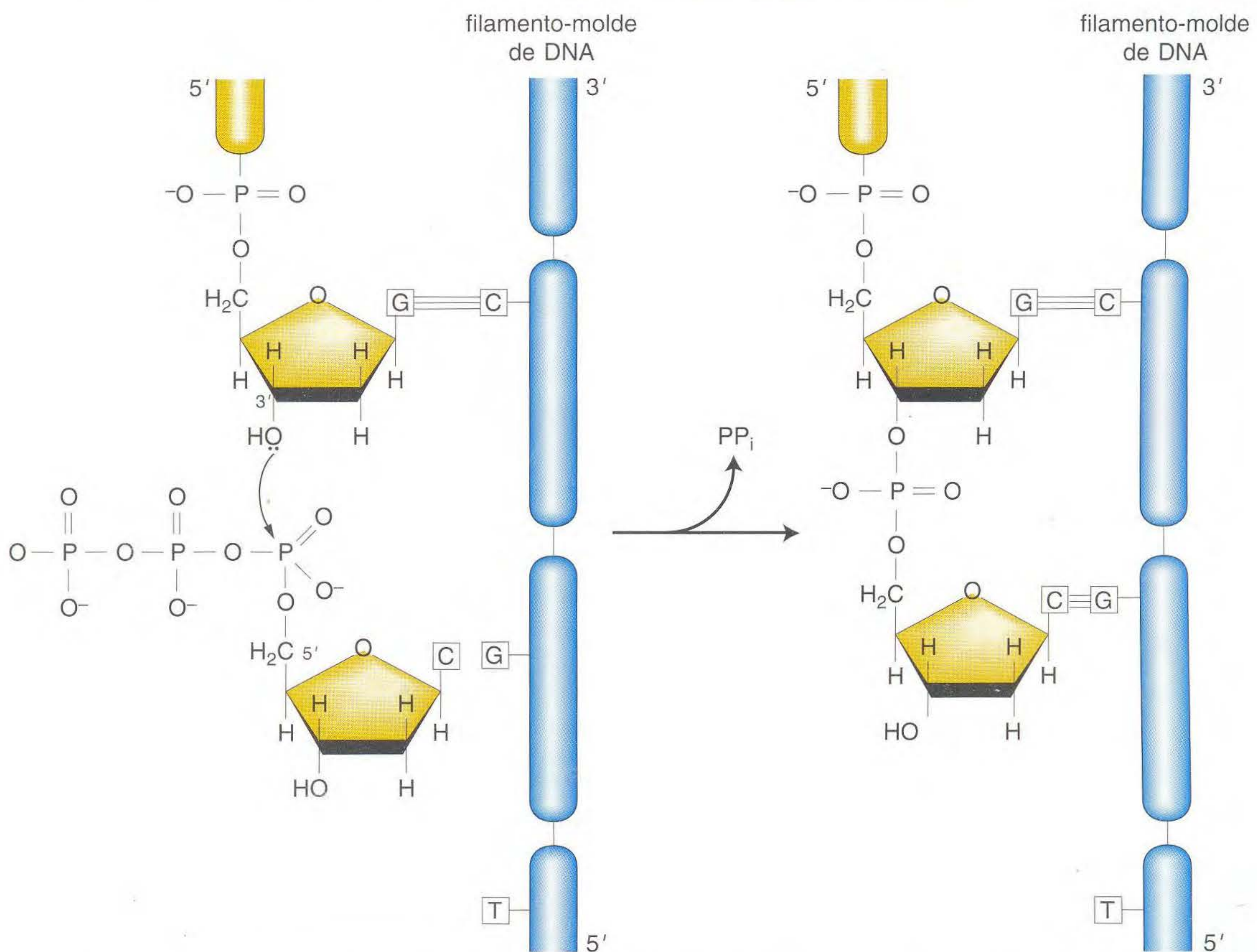


FIG. 7.15 A DNA polimerase catalisa a reação de alongamento da cadeia. A energia para a reação vem da quebra da ligação fosfato de alta energia do substrato trifosfato.

Existem hoje cinco DNA polimerases conhecidas em *E. coli*. A primeira enzima que Kornberg purificou hoje é chamada de DNA polimerase I, ou pol I. Essa enzima tem três atividades, que parecem estar localizadas em partes diferentes da molécula:

1. uma atividade polimerase, que catalisa o crescimento da cadeia no sentido 5' para 3';
2. uma atividade de exonuclease 3' para 5', que remove os pareamentos errados de bases; e
3. uma atividade de exonuclease 5' para 3', que degrada a dupla hélice de DNA.

Voltaremos ao significado das atividades das duas exonucleases mais adiante, neste mesmo capítulo.

Embora pol I tenha um papel na replicação do DNA (veja a seção seguinte), alguns cientistas suspeitam que não era possível para a maioria das sínteses de DNA, pois é muito lenta (~ 20 nucleotídeos/segundo) e muito abundante (~ 400 moléculas/célula) e porque ela se dissocia do DNA após a incorporação apenas de 20 a 50 nucleotídeos. Em 1969, John Cairns e Paula DeLucia esclareceram esse ponto quando demonstraram que uma linhagem de *E. coli* possuindo uma mutação no gene que codifica a DNA pol I ainda era capaz de crescer normalmente e replicar seu DNA. Eles concluíram que outra DNA polimerase, hoje chamada de pol III, catalisa a síntese de DNA na forquilha de replicação.

► 7.4 Visão Geral da Replicação do DNA

À medida que a DNA pol III avança, a dupla hélice é continuamente desenrolada na frente da enzima para expor mais os filamentos únicos de DNA que atuarão como moldes (Figura 7.16). A DNA pol III atua na **forquilha de replicação**, a zona onde a dupla hélice está se desenrolando. Entretanto, como a DNA polimerase sempre adiciona nucleotídeos na *ponta 3' cres-*

cente, apenas um dos dois filamentos de polaridade inversa pode servir como molde para a replicação no sentido da forquilha de replicação. Para esse filamento, a síntese pode ocorrer de modo contínuo no sentido da forquilha; o novo filamento sintetizado nesse molde é chamado de **filamento contínuo (leading)**.

A síntese do outro filamento também ocorre nas pontas crescentes 3', mas essa síntese está no sentido "errado", pois, para esse filamento, a síntese a direção 5'–3' está distante da forquilha de replicação (veja Fig. 7.16). Como veremos, a natureza da maquinaria de replicação requer que a síntese de ambos os filamentos ocorra na região da forquilha de replicação. Portanto, a síntese que se move afastando-se da forquilha de replicação não pode continuar por muito tempo. Ela deve ser em segmentos curtos: a polimerase sintetiza um segmento e, então, move-se para a ponta 5' do segmento, onde a forquilha crescente expõe um novo molde, e começa novamente o processo. Esses trechos curtos (1.000 a 2.000 nucleotídeos) de DNA recém-sintetizado são chamados de **fragmentos de Okazaki**.

Outro problema na replicação do DNA surge porque a DNA polimerase pode ampliar uma cadeia, mas não pode começar uma cadeia. Portanto, a síntese tanto do filamento leading quanto de cada fragmento de Okazaki deve ser iniciada por um **primer**, uma cadeia curta de nucleotídeos que se liga ao filamento-molde para formar um segmento de ácido nucleico dúplice. O primer na replicação do DNA pode ser visto na Figura 7.17. Os primers são sintetizados por um grupo de proteínas chamado **primossomo**, do qual um componente central é uma enzima chamada **primase**, um tipo de RNA polimerase. A primase sintetiza um trecho curto (8 a 12 nucleotídeos) de RNA complementar a uma região específica do cromossomo. No filamento leading, apenas um primer inicial é necessário, pois, após a ação do primer inicial, o filamento crescente de DNA serve como primer para a adição contínua. Entretanto, no filamento descontínuo (lagging), cada fragmento de Okazaki precisa de seu próprio primer. A cadeia de RNA que compõe o primer é então ampliada como uma cadeia de DNA pela DNA pol III.

Uma DNA polimerase diferente, a pol I, remove os primers de RNA e preenche os espaços resultantes com DNA. Como já mencionado, pol I é a enzima originalmente purificada por Kornberg. Outra enzima, a **DNA ligase**, junta a ponta 3' do DNA preenchedor do espaço à ponta 5' do fragmento de Okazaki posterior. O novo filamento assim formado é chamado de **filamento descontínuo (lagging)**. Uma DNA ligase une os pedaços quebrados do DNA, catalisando a formação de uma ligação fosfodiéster entre a extremidade 5'-fosfato de um fragmento e o grupo 3'-OH adjacente de outro fragmento.

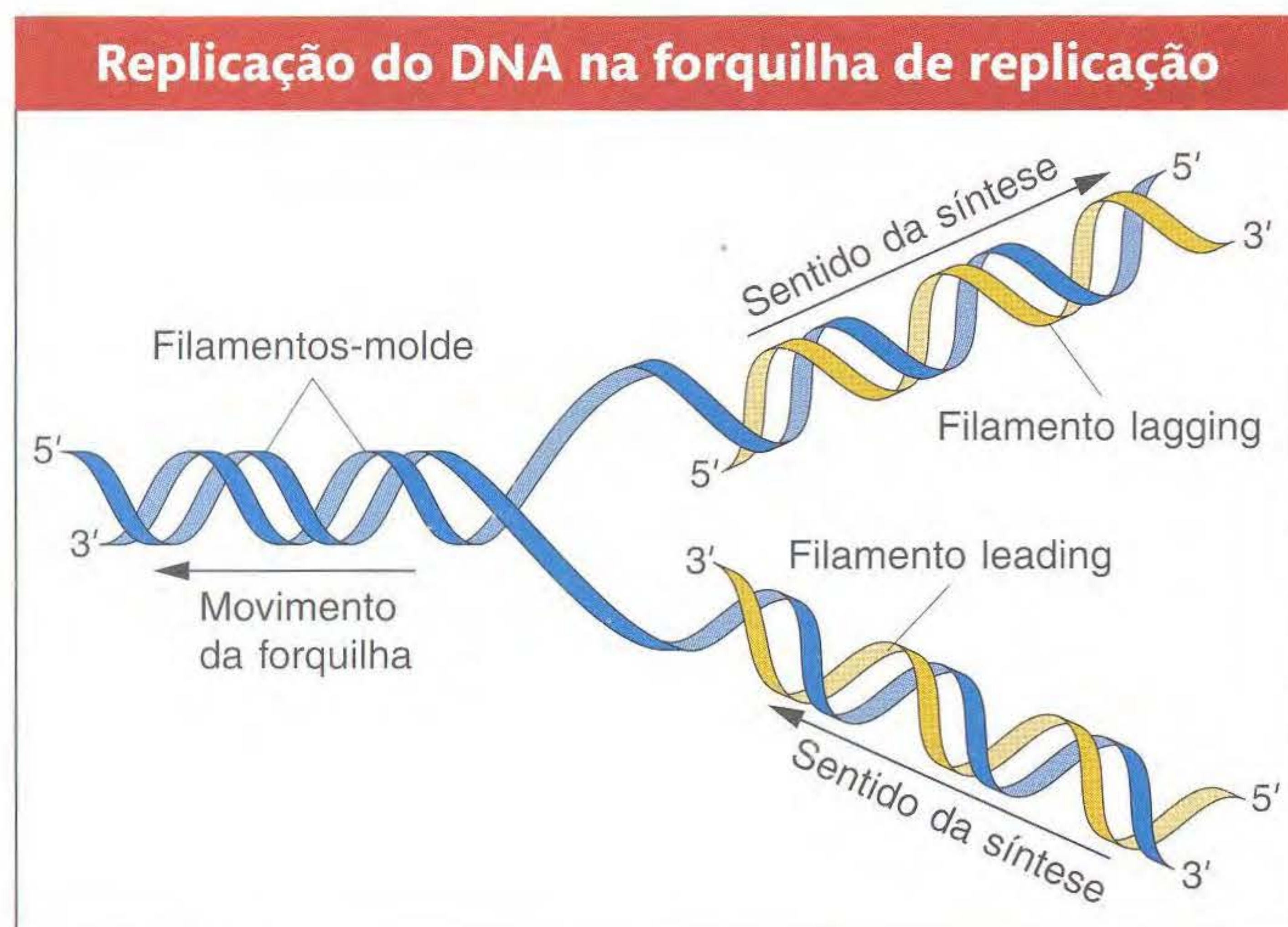
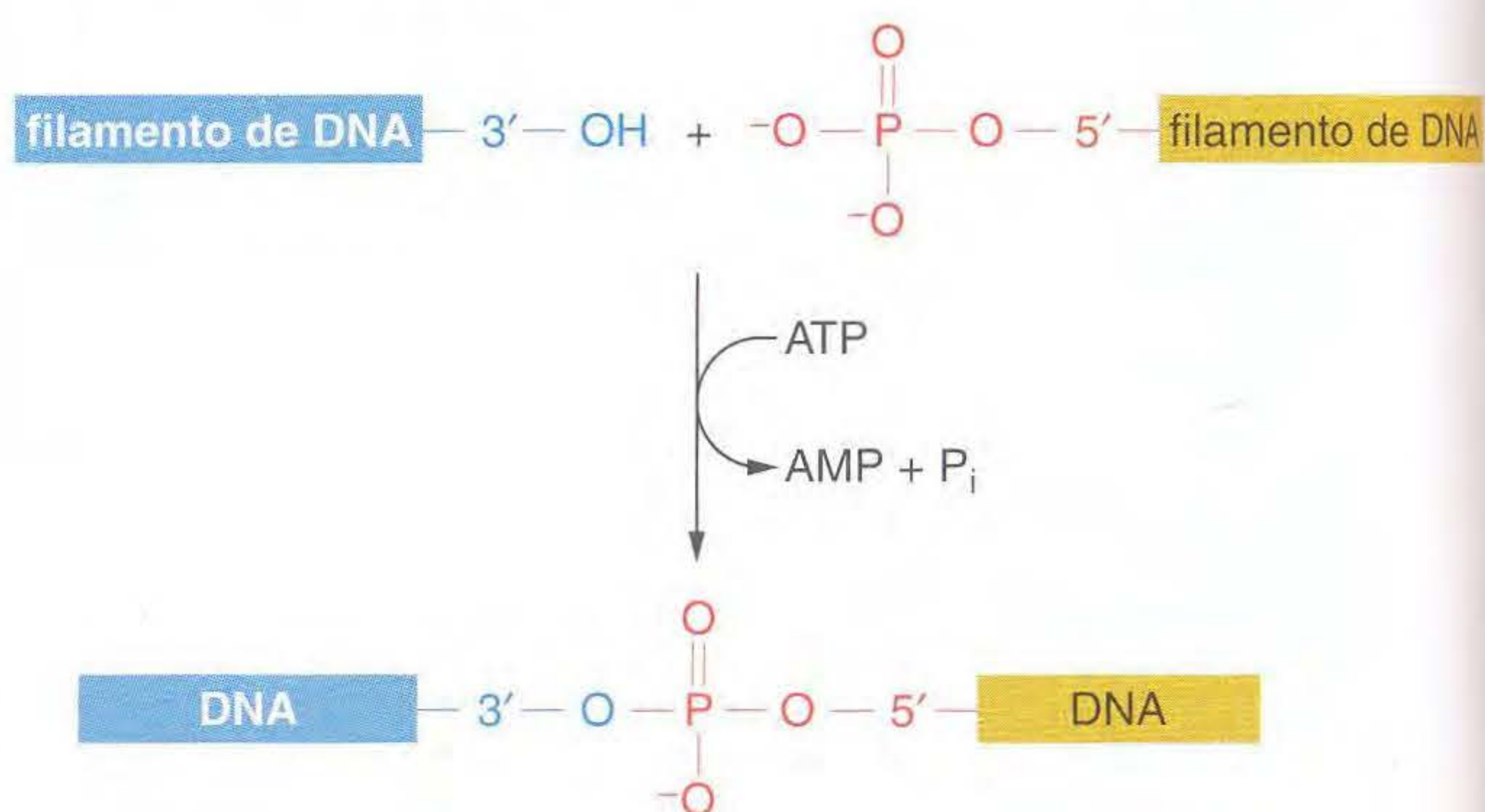


FIG. 7.16 A forquilha de replicação move-se na síntese de DNA à medida que a dupla hélice continuamente se deslicoidiza. A síntese do filamento contínuo (leading) pode continuar sem interrupção no sentido do movimento da forquilha de replicação, mas a síntese do filamento descontínuo (lagging) deve ocorrer no sentido oposto, afastando-se da forquilha de replicação.

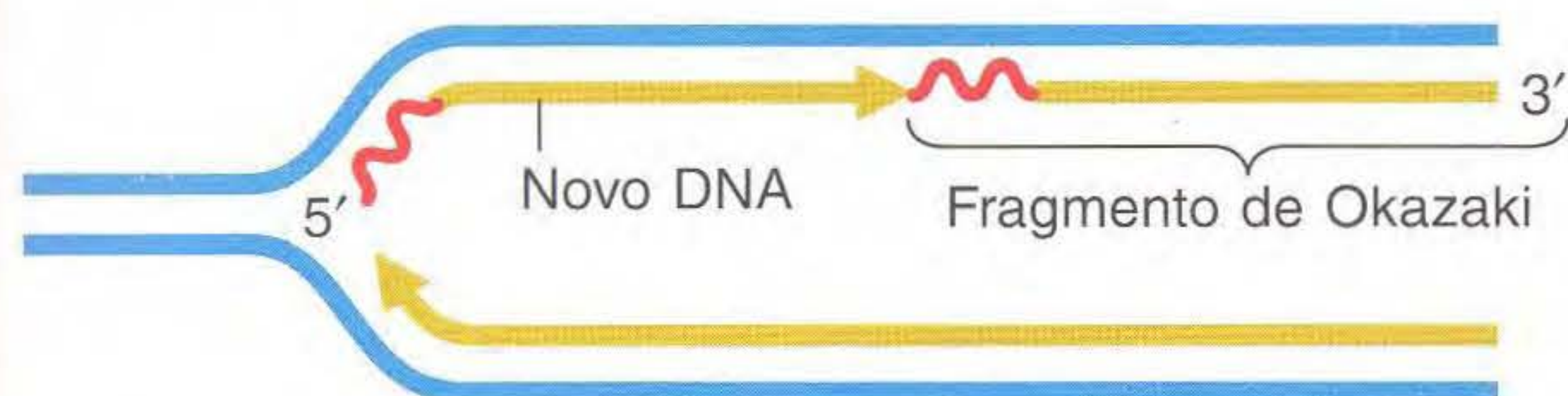


Síntese do filamento lagging

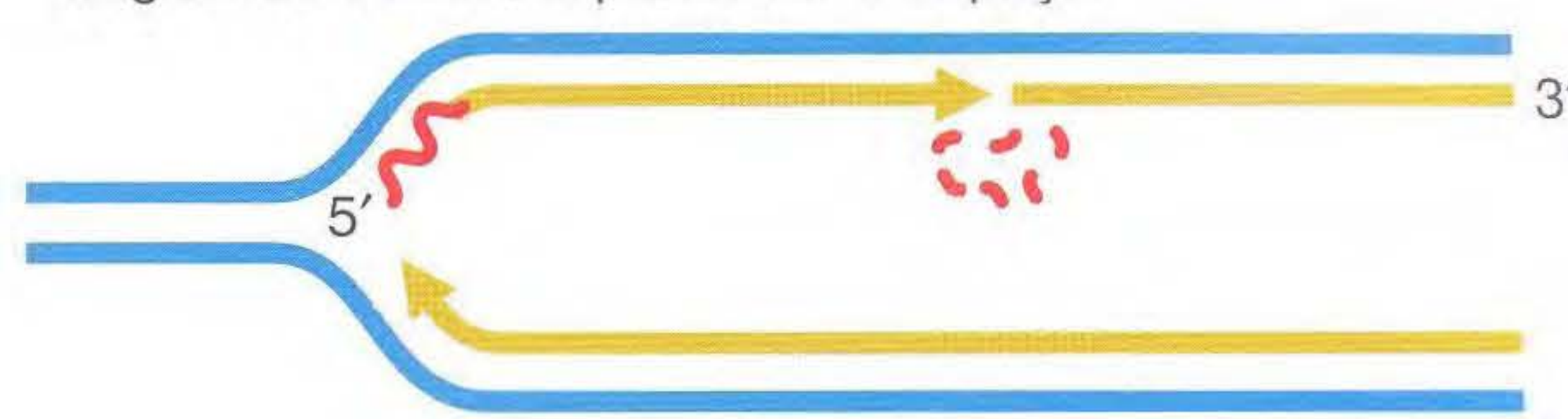
1. A primase sintetiza oligonucleotídeos curtos de RNA (primer) copiados do DNA.



2. A DNA polimerase III alonga os primers com novo DNA.



3. DNA polimerase I remove o RNA na ponta 5' do fragmento vizinho e preenche o espaço.



4. A DNA ligase conecta fragmentos adjacentes.

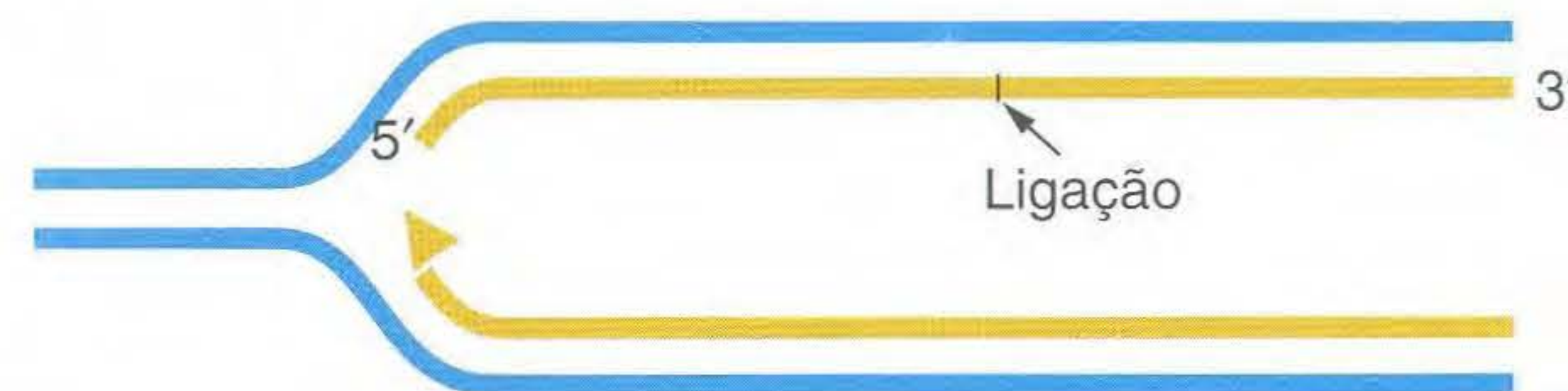


FIG. 7.17 Etapas na síntese do filamento descontínuo (lagging). A síntese de DNA ocorre pela síntese contínua do filamento leading e pela síntese descontínua do filamento lagging.

A característica da replicação do DNA é sua precisão, também chamada de fidelidade: no geral, é inserido menos de um erro por 10^{10} nucleotídeos. Parte do motivo da precisão da replicação do DNA é que tanto a DNA pol I quanto a DNA pol III possuem a atividade de exonuclease 3' para 5', que serve a uma função de “revisão”, excisando bases erradamente inseridas. As linhagens sem uma exonuclease 3' para 5' funcional têm uma taxa maior de mutação. Além disso, como a primase não tem uma função de revisão, o primer de RNA mais provavelmente contém erros do que o DNA. Para manter a alta fidelidade de replicação, os primers de RNA nas pontas dos fragmentos de Okazaki devem ser removidos e substituídos por DNA. Essa remoção e substituição é feita pela DNA pol I, que se associa à ponta 3' de um fragmento de Okazaki e catalisa a síntese de DNA para substituir o primer de RNA do fragmento de Okazaki adjacente. Antes da síntese, o primer de RNA é degradado pela atividade 5'-3' de exonuclease de pol I (veja Figura 7.17). O assunto do reparo do DNA será visto em detalhe no Capítulo 15.

A replicação do DNA ocorre na forquilha de replicação, onde a dupla hélice está se desenrolando e os dois filamentos estão se separando. A replicação do DNA ocorre continuamente no sentido da forquilha de replicação em deslicoidização no filamento de replicação contínua (leading). O DNA é sintetizado em segmentos curtos, afastando-se da forquilha de replicação, no filamento de replicação descontínuo. A DNA polimerase requer um primer, uma cadeia curta de nucleotídeos, posicionada para começar a síntese.

ser tão curto quanto 40 min. Portanto, seu genoma de cerca de 5 milhões de pares de bases deve ser copiado a uma velocidade de cerca de 2.000 nucleotídeos por segundo. Pelo experimento de Cairns, sabemos que a *E. coli* usa apenas duas forquilhas de replicação para copiar todo o seu genoma. Assim, cada forquilha deve ser capaz de mover-se a uma velocidade de 1.000 nucleotídeos por segundo. O que é incrível nesse processo todo de replicação do DNA é que ele não sacrifica a velocidade pela precisão. Como ela pode manter tanto a velocidade como a precisão, considerando a complexidade das reações na forquilha de replicação? A resposta é que a DNA polimerase é parte de um grande complexo “nucleoproteico” que coordena as atividades na forquilha de replicação. Esse complexo, chamado de **replissomo**, é um exemplo de uma “máquina molecular”. Você encontrará outros exemplos em capítulos posteriores. A descoberta de que a maioria das funções principais das células — replicação, transcrição e tradução, por exemplo — são feitas por grandes complexos multissubunitários mudou o modo pelo qual pensamos sobre a célula. Para começarmos a entender o motivo, vejamos o replissomo mais minuciosamente.

Alguns dos componentes que interagem no replissomo de *E. coli* são mostrados na Figura 7.18. Na forquilha de replicação, o cerne catalítico da DNA pol III é parte de um complexo muito maior, chamado de **holoenzima polimerase III (pol III)**, que consiste em dois cernes catalíticos e muitas **proteínas acessórias**. Um dos cernes catalíticos lida com a síntese do filamento contínuo, enquanto o outro lida com a síntese do filamento descontínuo. Algumas das proteínas acessórias (não visíveis na Figura 7.18) formam uma conexão que faz uma ponte entre dois cernes catalíticos, coordenando assim a síntese dos filamentos contínuo e descontínuo. O filamento descontínuo é mostrado fazendo uma alça de modo que o replissomo possa coordenar a síntese de ambos os filamentos e mover-se na direção da forquilha de replicação. Também mostrada é uma importante proteína acessória chamada de

7.5 O Replissomo: uma Marcante Máquina de Replicação

Outro marco da replicação do DNA é a velocidade. O tempo necessário para que a *E. coli* replique seu cromossomo pode

Proteínas funcionando na forquilha de replicação

Síntese dos filamentos leading e lagging

W.W. ANIMAÇÃO COMPUTADORIZADA

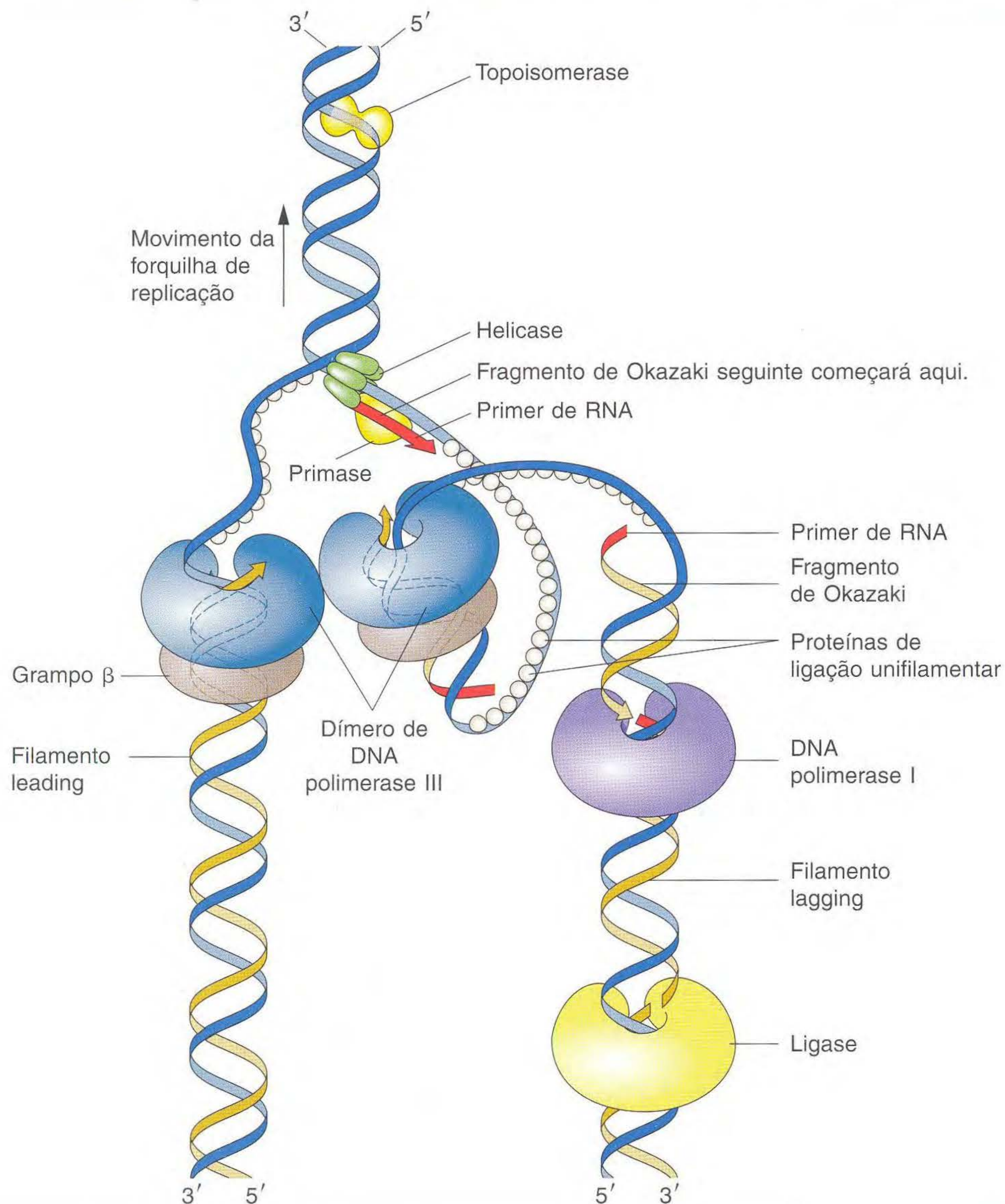


FIG. 7.18 O replissomo e as proteínas acessórias efetuam várias etapas na forquilha de replicação. A topoisomerase e a helicase desenrolam e abrem a dupla hélice na preparação para a replicação do DNA. Quando a dupla hélice está desenrolada, as proteínas de ligação a um só filamento impedem que a dupla hélice se reconstitua. A ilustração é uma representação do chamado modelo trombone (denominada por sua semelhança a um trombone devido à alça do filamento de replicação descontínua) mostrando como os dois cernes catalíticos do replissomo são vistos interagindo para coordenar os vários eventos da replicação dos filamentos leading e lagging. [Geoffrey Cooper, *The Cell*. Sinauer Associates, 2000.]

grampo β, que circunda o DNA como uma rosca (*donut*) e mantém a pol III ligada à molécula de DNA. Assim, pol III é transformada de uma enzima que pode adicionar apenas 10 nucleotídeos antes de sair do molde (chamada de **enzima distributiva**) em uma enzima que fica na forquilha em movimento e adiciona dezenas de milhares de nucleotídeos (uma **enzima processiva**). Em suma, pela ação de proteínas acessórias, a síntese de ambos os filamentos, contínuo e descontínuo, é rápida e altamente coordenada. Note também que a primase, a enzima que sintetiza o primer de RNA, não toca a proteína grampo. Portanto, a primase atua como uma

enzima distributiva; ela ajuda apenas alguns ribonucleotídeos antes de se dissociar do molde. Esse modo de ação faz sentido porque o primer só precisa ser longo o suficiente para formar um ponto inicial dúplice adequado para a DNA pol III.

Deselicoidizando a dupla hélice

Quando a dupla hélice foi proposta em 1953, uma grande objeção era que a replicação de tal estrutura iria requerer a deselicoidização da dupla hélice na forquilha de replicação e as quebras nas pontes de hidrogênio que juntam os fila-

mentos. Como o DNA poderia ser deselicoïdizado tão rapidamente e, mesmo que pudesse, não helicoidizar mais ainda o DNA atrás da forquilha e fazer um emaranhado? Sabemos hoje que o replissomo contém duas classes de proteínas que abrem a hélice e impedem a maior helicoidização: elas são as **helicases** e as **topoisomerases**, respectivamente. As helicases são enzimas que rompem as pontes de hidrogênio que mantêm os dois filamentos da dupla hélice juntos. Como a proteína grampo, a helicase ajusta-se como uma rosca ao redor do DNA; dessa posição, ela rapidamente desenrola a dupla hélice à frente da síntese de DNA. O DNA desenrolado é estabilizado por **proteínas de ligação unifilamentares (SSB)**, que se ligam ao DNA unifilamentar e impedem a reestruturação do dúplice.

O DNA circular pode ser torcido e helicoidizado, como as voltas extras que podem ser introduzidas em um elástico. O desenrolamento da forquilha de replicação pelas helicases causa torções extras em outras regiões, e hélices chamadas

super-hélices formam-se para liberar a tensão da torção extra. Ambos os giros e super-hélices devem ser removidos para permitir que a replicação continue. Essa super-helicoidização pode ser criada ou relaxada por enzimas chamadas topoisomerases, um exemplo das quais é a DNA girase (Figura 7.19). As topoisomerases relaxam o DNA super-helicoidizado quebrando um único filamento de DNA ou ambos os filamentos, o que permite que o DNA gire para uma molécula relaxada. A topoisomerase termina reunindo os filamentos da molécula de DNA agora relaxada.

Uma máquina molecular chamada replissomo faz a síntese de DNA. Ela inclui duas unidades de DNA polimerase para fazer a síntese em cada filamento e coordenar a atividade de proteínas acessórias necessárias para o priming, deselicoïdização da dupla hélice e estabilização dos filamentos isolados.

Montagem do replissomo: início da replicação

A montagem do replissomo é um processo ordenado que começa em locais precisos no cromossomo (chamados **origens**) e ocorre apenas em certas épocas da vida da célula. A replicação de *E. coli* começa em uma origem fixa (chamada *oriC*) e, então, continua em ambos os sentidos (com as forquilhas movendo-se em ambas as pontas, como mostrado na Figura 7.14) até que as forquilhas se juntem. A Figura 7.20 mostra o processo. A primeira etapa na montagem do replissomo é a ligação de uma proteína chamada DnaA a uma sequência específica de 13 pares de bases (pb) (chamada "DnaA box") que é repetida cinco vezes em *oriC*. Em resposta à ligação do DnaA, a origem é a deselicoïdização em um grupo de nucleotídeos A e T. Lembre que os pares de bases AT são mantidos juntos apenas com duas pontes de hidrogênio, enquanto os pares de bases GC são mantidos juntos por três. Assim, é mais fácil separar (dissociar) a dupla hélice em trechos do DNA que são enriquecidos em bases A e T.

Após começar a deselicoïdização, as proteínas DnaA ligam-se a regiões unifilamentares recém-deselicoïdizadas. Com DnaA revestindo a origem, as duas helicases (proteína DnaB) agora se ligam e deslizam em uma direção 5' para 3' para começar a abrir a hélice na forquilha de replicação. A primase e a holoenzima DNA pol III agora são recrutadas para a forquilha de replicação por interações proteína-proteína, e começa a síntese do DNA. Você pode imaginar por que o DnaA não está presente na Figura 7.18 (a máquina de replissomo). A resposta é que, embora seja necessário para a montagem do replissomo, não é parte da maquinaria de replicação. Sua tarefa é levar o replissomo para o lugar correto no cromossomo circular, para o início da replicação.

7.6 Replicação em Organismos Eucarióticos

A replicação do DNA, tanto em procariontes quanto em eucariontes, usa um mecanismo semiconservativo e emprega a síntese de filamentos de replicação contínua e descontínua.

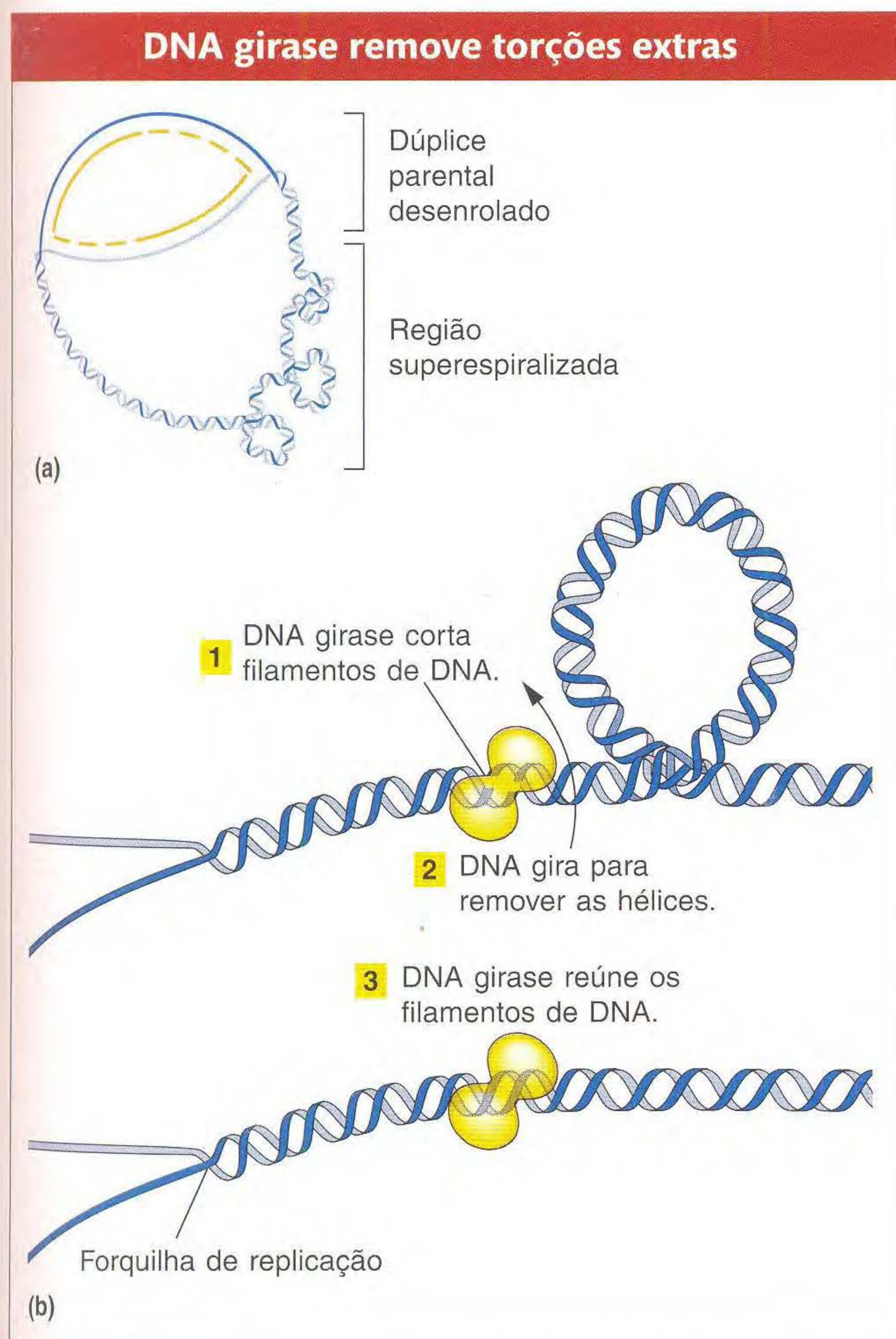


FIG. 7.19 A DNA girase, uma topoisomerase, remove giros extras durante a replicação. (a) As regiões extras torcidas (positivamente super-helicoidizadas) acumulam-se à frente da forquilha à medida que os filamentos parentais se separam para replicação. (b) Uma topoisomerase tal como a DNA girase remove essas regiões, cortando os filamentos de DNA, permitindo que eles girem e, então, reagrupando os filamentos. [(a) De A. Kornberg and T. A. Baker, *DNA Replication*, 2nd ed. Copyright 1992 by W.H. Freeman and Company.]

Iniciação procariótica da replicação

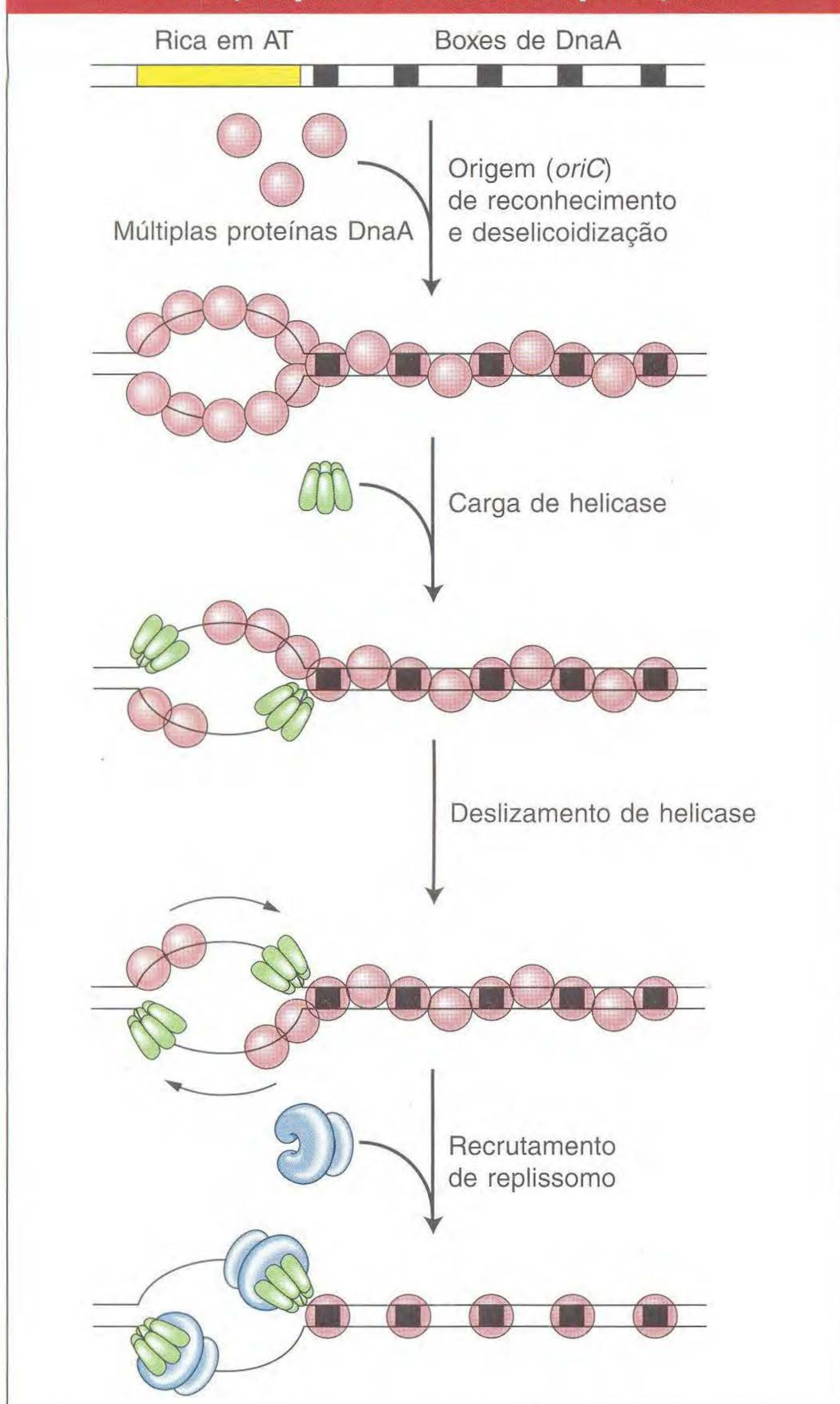


FIG. 7.20 A síntese de DNA é iniciada nas origens da replicação em procariontes. As proteínas ligam-se à origem (*oriC*), onde separam os dois filamentos da dupla hélice e recrutam componentes do replissomo para as duas forquilhas de replicação.

Por esse motivo, não deve ser surpresa que os componentes do replissomo procariótico e os do replissomo eucariótico sejam muito similares. Entretanto, à medida que os organismos aumentam em complexidade, o número de componentes do replissomo também aumenta.

O replissomo eucariótico

Hoje sabemos que existem 13 componentes do replissomo e, pelo menos, 27 nos replissomos de leveduras e mamíferos. Um motivo para a complexidade adicional do replissomo eucariótico é a maior complexidade do molde eucariótico. Lembre que, ao contrário do cromossomo bacteriano, os cromossomos eucarióticos existem no núcleo como cromatina. Como descrito no Capítulo 2, a unidade básica da cromatina é o **nucleossomo**, que consiste em DNA enrolado ao redor de

proteínas histonas. Assim, o replissomo não tem só que copiar os filamentos parentais, mas também desmontar os nucleossomos nos filamentos parentais e os reunir nas moléculas-filhas. Essa manobra é feita distribuindo-se aleatoriamente as histonas antigas (dos nucleossomos existentes) para as moléculas-filhas e levando novas histonas em associação a uma proteína chamada fator 1 de montagem de cromatina (CAF-1) para o replissomo. CAF-1 liga-se a histonas e as direciona para a forquilha de replicação, onde podem ser montadas junto com DNA recém-sintetizado. CAF-1 e sua carga de histonas chegam na forquilha de replicação ligando-se à versão eucariótica do grampo β , chamado **antígeno nuclear de proliferação celular (PCNA)** (Figura 7.21).

O replissomo eucariótico faz todas as funções do replissomo procariótico; em adição, ele pode desmontar e remontar os complexos proteína-DNA chamados nucleossomos.

Origens eucarióticas da replicação

Bactérias tais como *E. coli* geralmente completam o ciclo de replicação-divisão em 20 a 40 min, mas, nos eucariontes, o ciclo pode variar de 1,4 h em leveduras a 24 h em células animais cultivadas, e podem durar de 100 a 200 h em algumas células. Os eucariontes têm que resolver o problema de coordenar a replicação de mais de um cromossomo, bem como o problema de replicar a própria estrutura complexa do cromossomo.

Para compreender as origens da replicação eucariótica, primeiro voltaremos nossa atenção para o eucarionte simples, a levedura. Muitas proteínas eucarióticas com papéis nas origens de replicação foram primeiro identificadas em leveduras devido à facilidade da análise genética em pesquisas de leveduras (veja o boxe Organismo-modelo *Levedura* no Capítulo 11). As origens de replicação em leveduras são muito parecidas com *oriC* em *E. coli*. As origens de 100 a 200 pb têm uma sequência de DNA conservada que inclui uma região rica em AT que se dissocia quando uma proteína iniciadora se liga a sítios adjacentes de ligação. Ao contrário dos cromossomos procarióticos, cada cromossomo eucariótico tem muitas origens de replicação para replicar rapidamente os genomas eucarióticos maiores. Aproximadamente 400 origens de replicação são dispersas pelos 16 cromossomos de levedura, e são estimadas cerca de milhares de forquilhas de replicação nos 23 cromossomos humanos. Assim, nos eucariontes, a replicação continua em ambas as direções a partir de múltiplos pontos de origem (Figura 7.22). As duplas hélices que estão sendo produzidas em cada origem de replicação alongam-se e, ocasionalmente, juntam-se. Quando a replicação dos dois filamentos é completa, resultam duas **moléculas-filhas** idênticas de DNA.

Onde e quando ocorre a replicação são cuidadosamente controlados por uma montagem ordenada do replissomo em um local preciso, chamado de origem. A replicação continua em ambas as direções de uma única origem do cromossomo circular procariótico. A replicação continua em ambas as direções a partir de centenas ou milhares de origens em cada um dos cromossomos eucarióticos lineares.

Montagem de nucleossomos durante a replicação do DNA

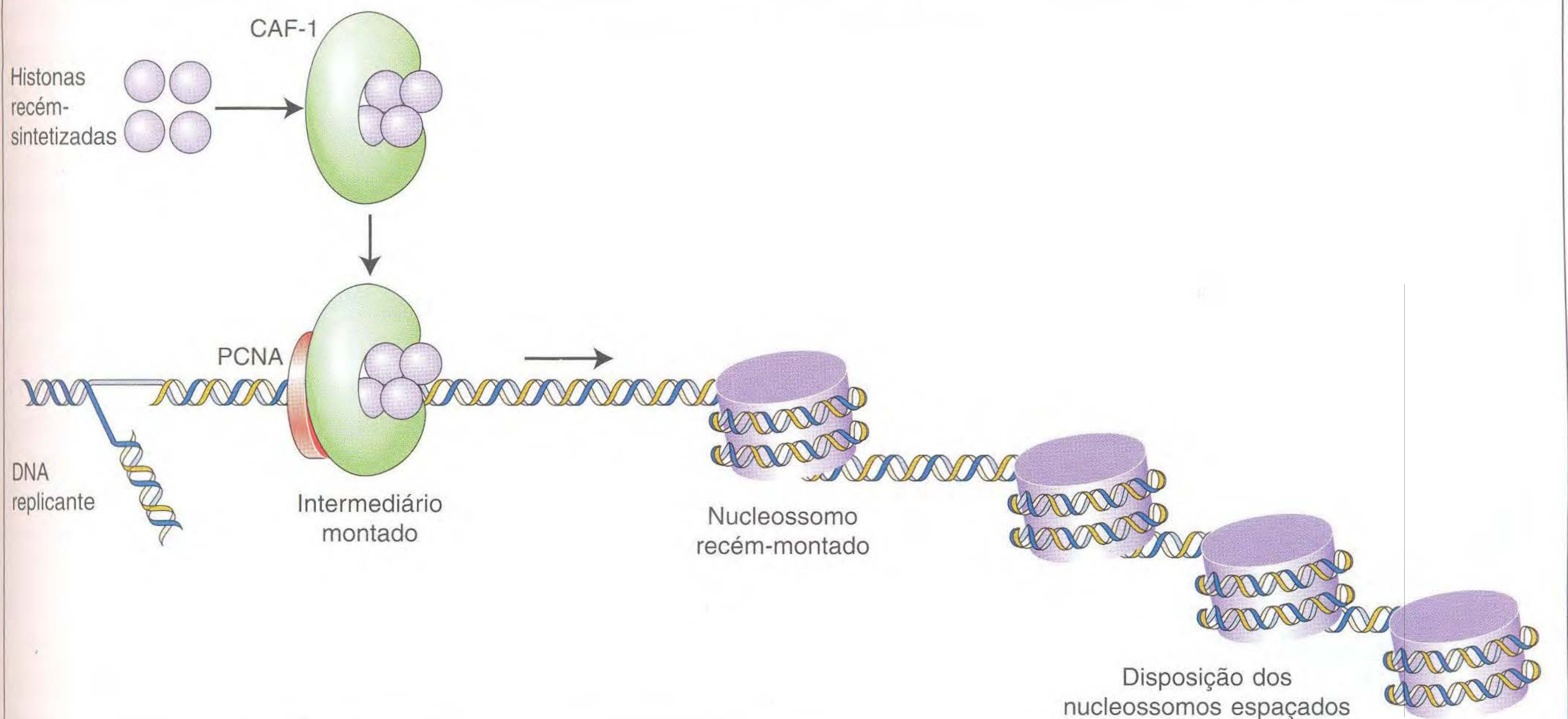


FIG. 7.21 A proteína CAF-1 (fator 1 de montagem de cromatina) leva histonas para a forquilha de replicação, onde são montadas para formar nucleossomos. PCNA, antígeno nuclear de proliferação celular. Por simplificação, a montagem do nucleossomo é mostrada apenas em um dos filamentos replicados de DNA.

A replicação do DNA ocorre em dois sentidos

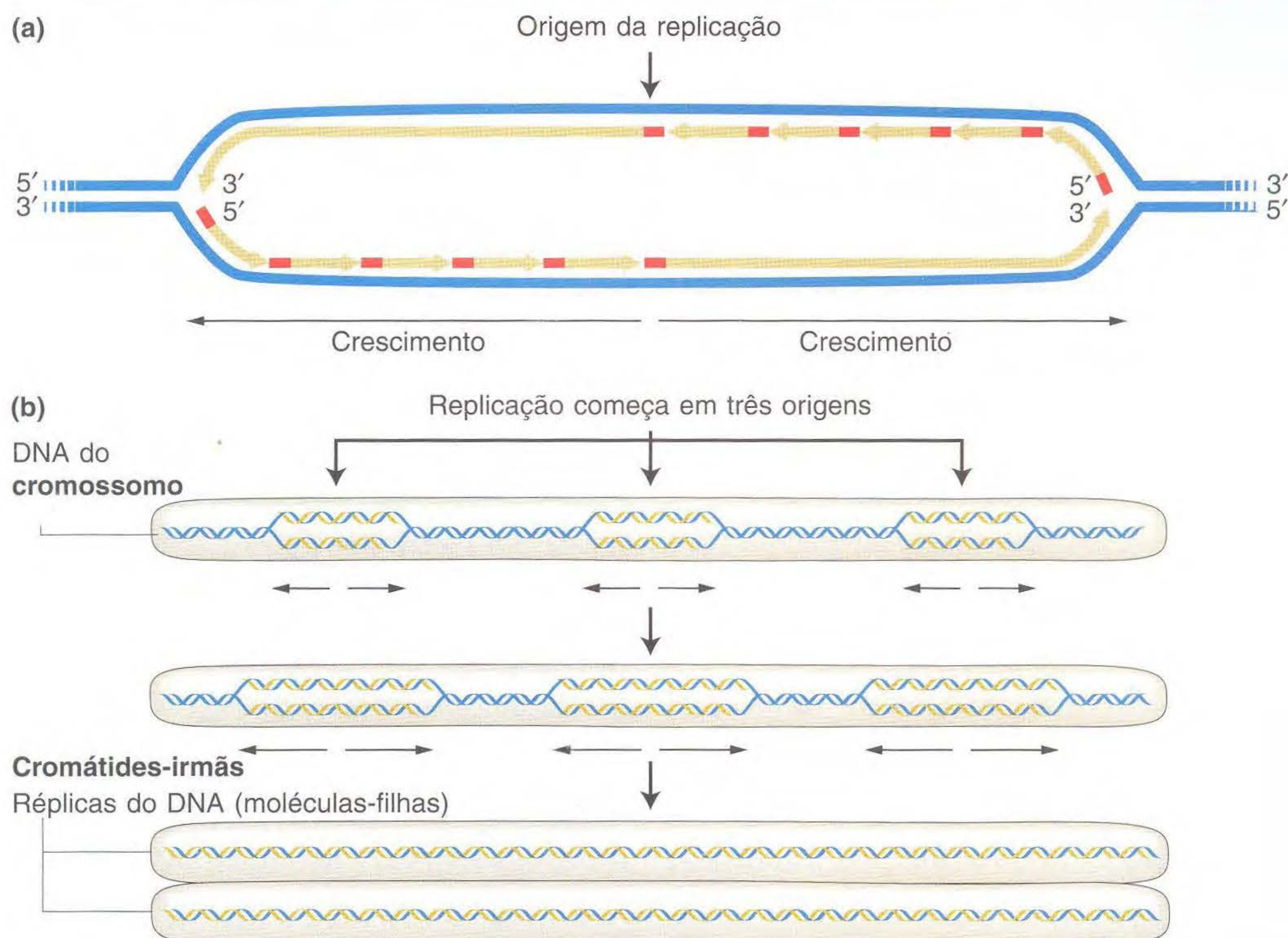


FIG. 7.22 A replicação do DNA ocorre em ambos os sentidos, a partir de uma origem de replicação. As setas pretas indicam o sentido do crescimento das moléculas-filhas de DNA. (a) Começando na origem, as DNA polimerases movem-se para fora em ambos os sentidos. As longas setas laranjas representam os filamentos de replicação contínua (leading), e as pequenas setas laranjas representam os filamentos descontínuos (lagging). (b) Como a replicação ocorre ao nível cromossômico. São mostradas três origens de replicação nesse exemplo.



Replicação do DNA e o ciclo celular de levedura

A síntese de DNA ocorre na fase S (síntese) do ciclo celular eucariótico (Figura 7.23). Como o início da síntese de DNA é limitada a esse único estágio? Em leveduras, o método de controle é ligar a montagem do replissomo ao ciclo celular. A Figura 7.24 mostra o processo. Em leveduras, são necessárias três proteínas para começar a montagem do replissomo. A origem do complexo de reconhecimento (ORC) primeiro liga-se a seqüências nas origens de leveduras, como a proteína DnaA faz em *E. coli*. A presença de ORC na origem serve para recrutar duas outras proteínas, Cdc6 e Cdt1. Ambas as proteínas mais ORC então recrutam a helicase replicativa, chamada de complexo MCM, e outros componentes do replissomo.

A replicação é ligada ao ciclo celular pela disponibilidade de Cdc6 e Cdt1. Em levedura, essas proteínas são sintetizadas durante o final da mitose e o intervalo 1 (G1) e são destruídas por proteólise após a síntese ter começado. Desse modo, o replissomo pode ser montado apenas antes da fase S. Quando começa a replicação, novos replissomos não podem formar-se nas origens, pois Cdc6 e Cdt1 são degradados durante a fase S e não estão mais disponíveis.

Origens de replicação em eucariontes superiores

Como já foi dito, a maioria das cerca de 400 origens de replicação em leveduras são compostas de seqüências similares de DNA (100 a 200 pb de tamanho) que são reconhecidas

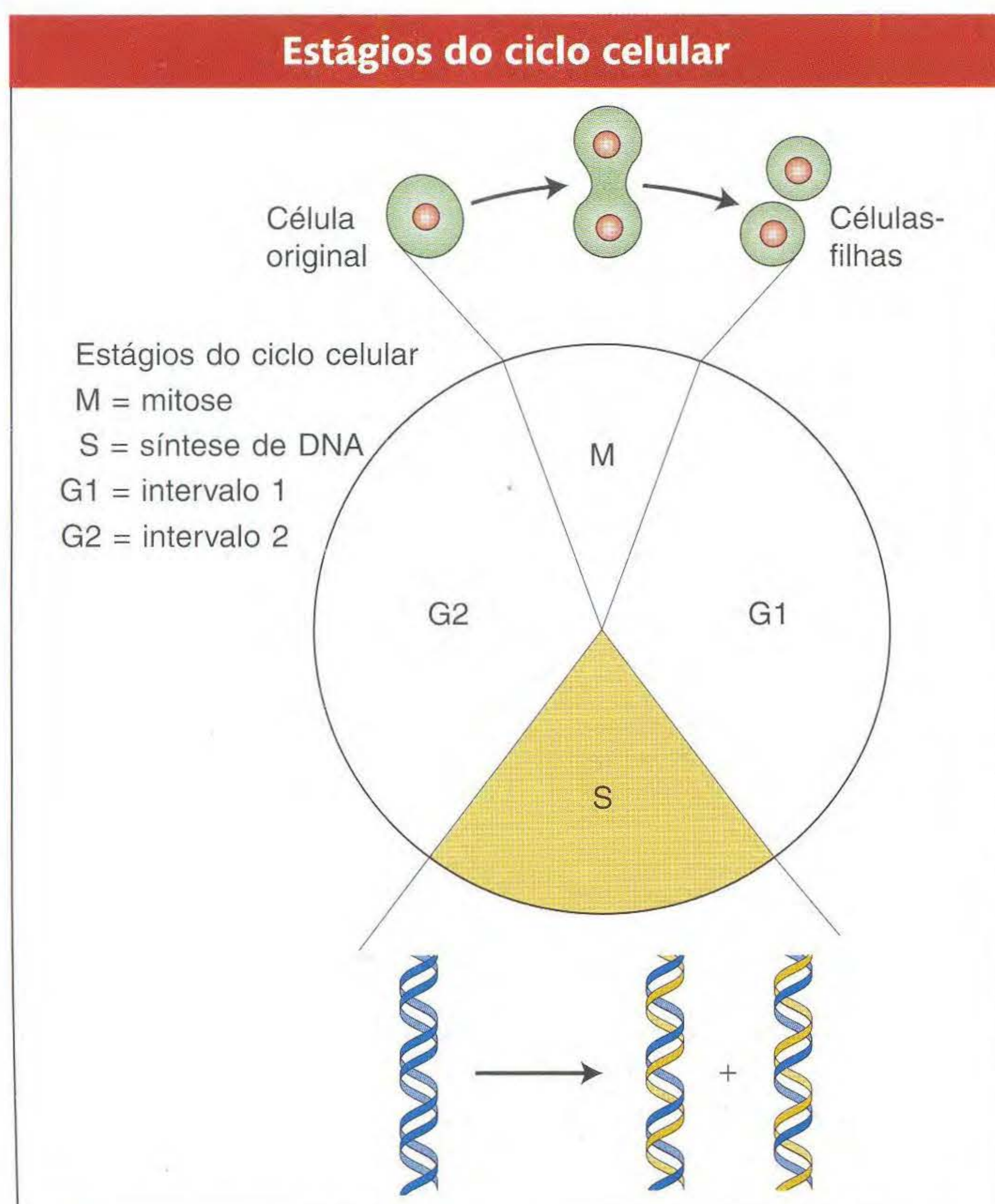


FIG. 7.23 O DNA é replicado durante a fase S do ciclo celular.

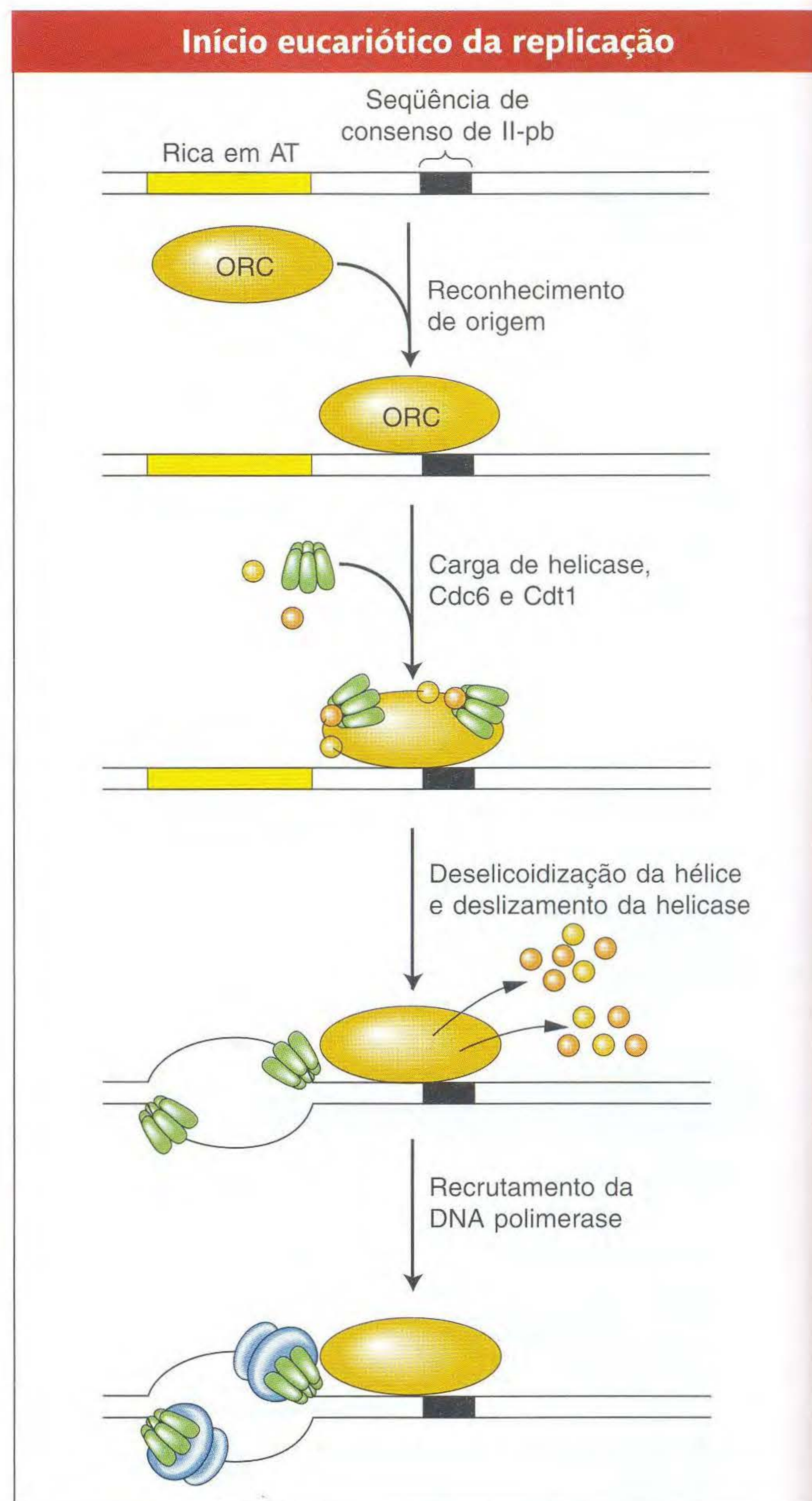


FIG. 7.24 Esse exemplo de levedura mostra o início da síntese de DNA em uma origem de replicação em um eucarionte. Como na iniciação procariótica (veja Figura 7.20), as proteínas do complexo de origem de reconhecimento (ORC) ligam-se à origem, onde separam os dois filamentos da dupla hélice e recrutam componentes do replissomo nas duas forquilha de replicação. A replicação é ligada ao ciclo celular pela disponibilidade das duas proteínas: Cdc6 e Cdt1.

por subunidades ORC. Curiosamente, embora todos os eucariontes caracterizados tenham proteínas ORC similares, as origens de replicação nos eucariontes superiores são maiores, possivelmente dezenas de milhares ou centenas de milhares de nucleotídeos. Significativamente, eles têm semelhanças de seqüência limitadas. Assim, embora as ORC de leveduras reconheçam seqüências específicas de DNA nos cromossomos de levedura, o que as ORC correlatas de eucariontes superiores reconhecem ainda não está claro, mas a característica reconhecida é provavelmente não uma seqüência específica de DNA. O que essa incerteza significa em ter-

mos práticos é que é muito difícil isolar origens de humanos e outros eucariontes superiores porque os cientistas não podem usar uma sequência de DNA isolada de uma origem humana, por exemplo, para fazer uma pesquisa em computador de toda a sequência do genoma humano para encontrar outras origens.

Se as ORC de eucariontes superiores não interagem com uma sequência específica dispersa pelos cromossomos, então como elas encontram as origens de replicação? Essas ORC são tidas como interagindo indiretamente com as origens por associação com outros complexos proteicos que são ligados a cromossomos. Tal mecanismo de reconhecimento pode ter evoluído de modo que os eucariontes superiores possam regular a época da replicação do DNA durante a fase S (veja o Capítulo 11 para mais sobre a eucromatina e heterocromatina). As regiões ricas em genes do cromossomo (a eucromatina) foram conhecidas por algum tempo como replicando-se cedo na fase S, enquanto as regiões pobres em genes, incluindo a heterocromatina densamente compactada, replicam-se tardiamente na fase S. A replicação do DNA não poderia ser regulada por região se as ORC estivessem ligadas a sequências relacionadas a sequências dispersas pelos cromossomos. Em vez disso, as ORC podem, por exemplo, ter uma afinidade maior por origens em cromatina aberta e ligar-se a essas origens primeiro e, então, ligar-se à cromatina condensada apenas após as regiões ricas em genes terem se replicado.

A origem de replicação de levedura, como a origem em procariontes, contém uma sequência conservada de DNA que é reconhecida pela ORC e outras proteínas necessárias para a montagem do replissomo. Em contraste, as origens de eucariontes superiores são difíceis de isolar e estudar porque são longas e complexas, e não contêm uma sequência conservada de DNA.

7.7 Telômeros e Telomerase: Término da Replicação

A replicação da molécula linear de DNA em um cromossomo eucariótico ocorre em ambas as direções a partir de várias origens de replicação, como mostrado na Figura 7.22. Esse processo replica a maioria do DNA cromossômico, mas há um problema inerente em replicar as duas pontas das moléculas lineares de DNA, as regiões chamadas de **telômeros**. A síntese contínua do filamento leading pode ocorrer até a ponta do molde. Entretanto, a síntese do filamento descontínuo (lagging) requer primers à frente do processo; logo, quando o último primer é removido, resta uma ponta unifilamentar em uma molécula-filha de DNA (Figura 7.25). Se o cromossomo-filho com essa molécula de DNA fosse replicado novamente, o filamento faltando sequências na ponta seria uma molécula bifilamentar encurtada após a replicação. A cada ciclo subsequente de replicação, o telômero continuaria a se encurtar, até que informações codificantes essenciais fossem perdidas.

O problema da replicação nas pontas dos cromossomos

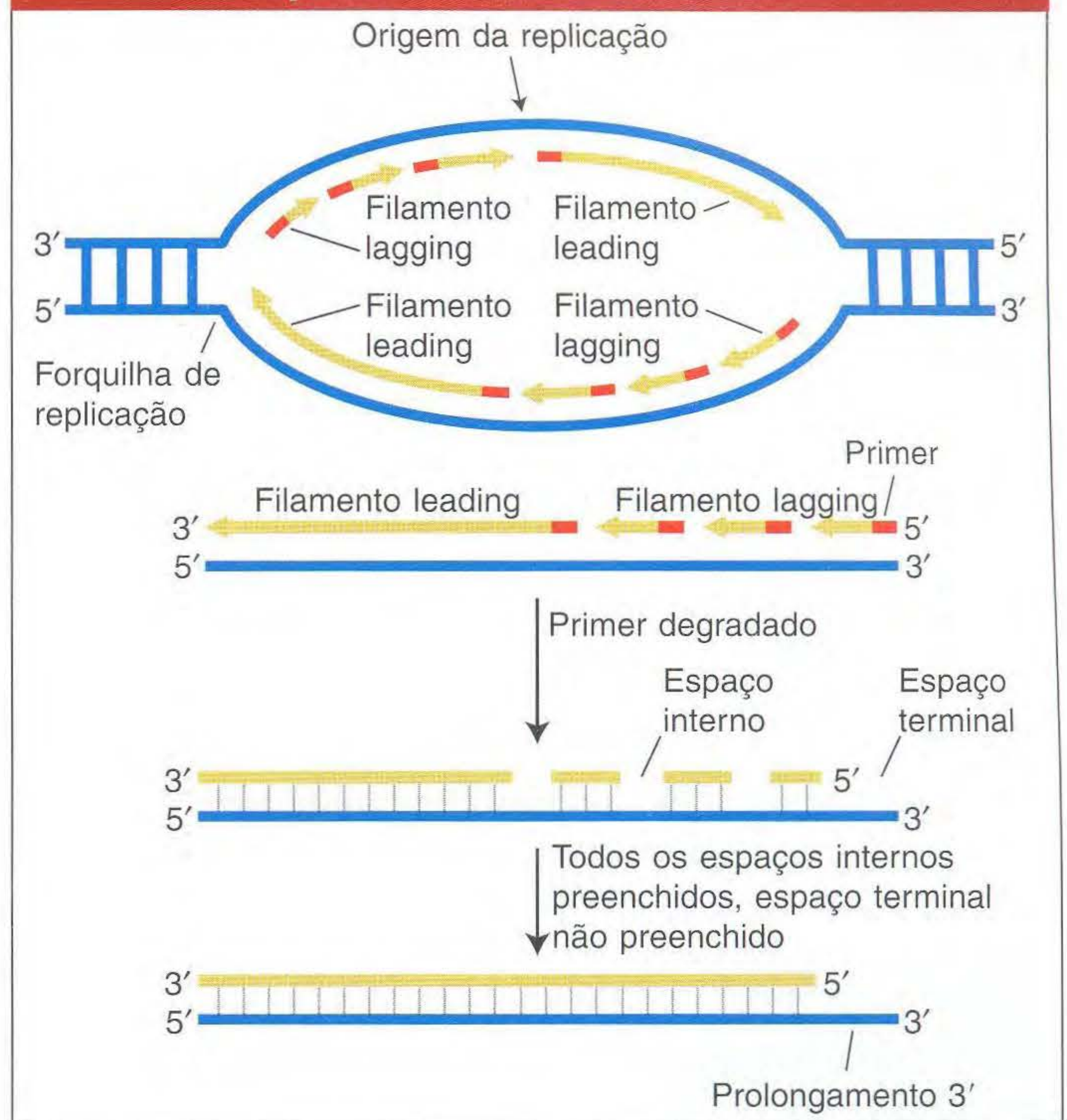


FIG. 7.25 (Em cima) A replicação de cada fragmento de Okazaki no filamento lagging começa com a inserção de um primer. (Embaixo) O destino do filamento inferior na bolha de transcrição. Quando o primer do último fragmento de Okazaki do filamento lagging é removido, não há modo para preencher o espaço por replicação convencional. Um cromossomo encurtado resultaria quando o cromossomo contendo o espaço fosse replicado.

As células desenvolveram um sistema especializado para evitar essa perda. Um componente do sistema lida com a adição de múltiplas cópias de uma sequência simples não-codificante ao DNA das pontas dos cromossomos. A descoberta de que as pontas dos cromossomos são feitas de sequências repetidas em tandem foi feita em 1978 por Elizabeth Blackburn e Joe Gall, que estavam estudando o DNA no macronúcleo incomum do ciliado unicelular *Tetrahymena*. Como outros ciliados, *Tetrahymena* tem um micronúcleo convencional e um macronúcleo incomum no qual os cromossomos são fragmentados em milhares de pedaços do tamanho de genes com novas pontas adicionadas a cada pedaço. Blackburn e Gall foram capazes de isolar os fragmentos contendo os genes para RNA ribossômico (fragmentos chamados de rDNA, veja o Capítulo 9 para mais sobre os ribossomos) usando centrifugação em gradiente de CsCl, a técnica desenvolvida por Meselson e Stahl para isolar DNA recém-replicado de *E. coli* (veja anteriormente). As pontas dos fragmentos de rDNA contêm disposições em tandem da sequência TTGGGG. Hoje sabemos que, praticamente, todos os eucariontes têm curtas repetições em tandem em suas pontas de cromossomos; entretanto, a sequência não é exatamente a mesma. Os cromossomos humanos, por exemplo, terminam em cerca de 10 a 15 kb de repetições em tandem da sequência TTAGGG.

Como essas repetições impedem a perda de DNA dos telômeros após cada rodada de replicação do DNA? Elizabeth Blackburn e Carol Greider formularam a hipótese de que as repetições foram adicionadas às pontas do cromossomo por uma enzima. Trabalhando novamente com extratos de macronúcleo de *Tetrahymena* (com seus cerca de 40.000 telômeros), eles identificaram uma enzima, que chamaram de **telomerase**, que adiciona repetições curtas às pontas 3' das moléculas de DNA. Curiosamente, a proteína telomerase leva uma pequena molécula de RNA, parte da qual atua como um molde para a polimerização da unidade repetida telomérica. Em todos os vertebrados, incluindo os humanos, a sequência de RNA 3'-AAUCCC-5' atua como um molde para a unidade repetida 5'-TTAGGG-3' por um mecanismo mostrado na Figura 7.26. Resumidamente, a RNA telomerase primeiro se helicoidiza ao prolongamento 3' do DNA, que é então ampliado com o uso de dois componentes da telomerase: o pequeno RNA (como molde) e a proteína (como atividade de polimerase). Após a adição de alguns nucleotídeos ao prolongamento 3', a RNA polimerase move-se ao longo do DNA de modo que a ponta 3' possa ser mais estendida por sua atividade de polimerase. A ponta 3' continua a ser ampliada pelo movimento repetido da RNA polimerase. A primase e a DNA polimerase então usam o prolongamento 3' como molde para preencher o final do outro filamento de DNA.

Telômeros, câncer e envelhecimento

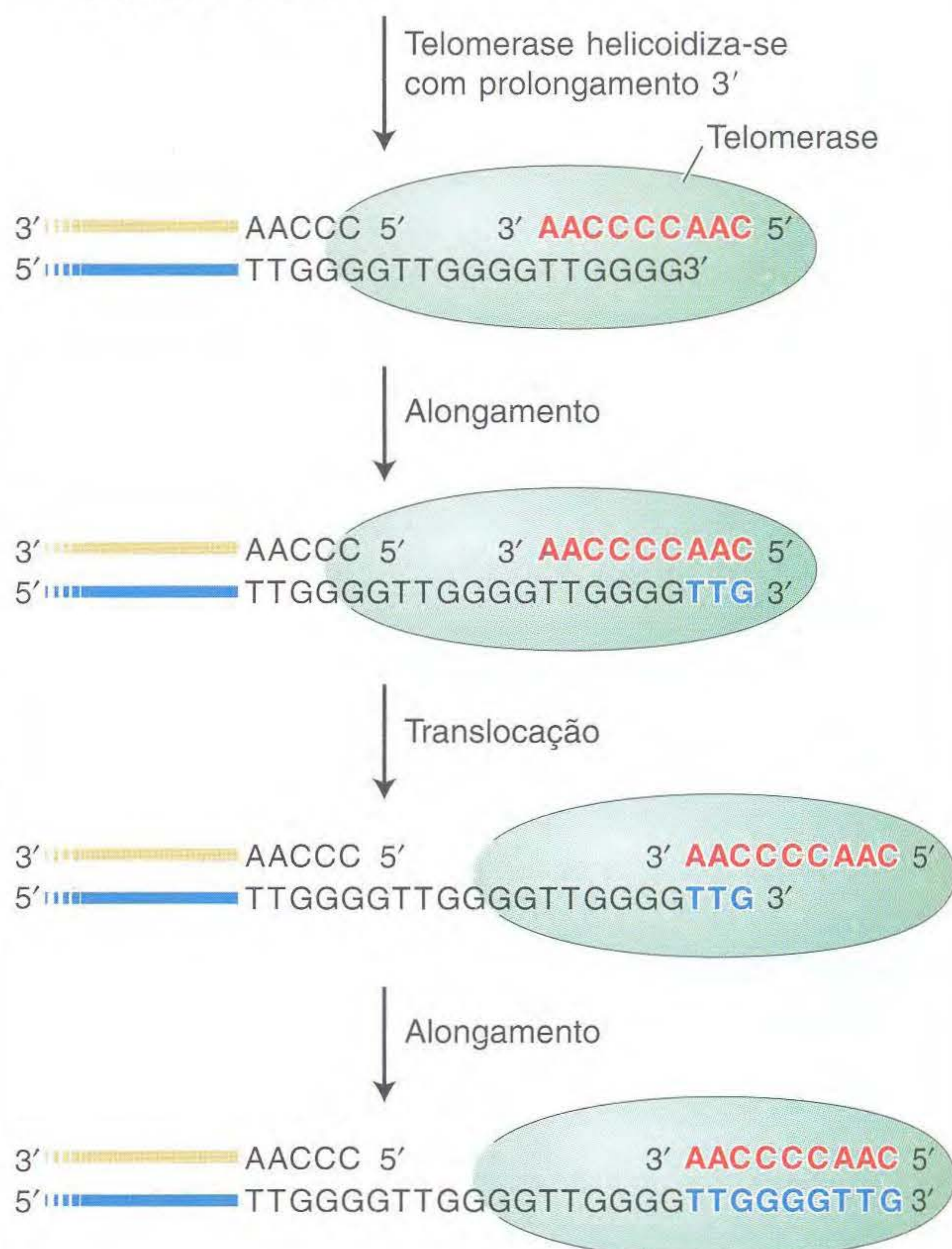


Além de impedir a erosão do material genético após cada rodada de replicação, os telômeros preservam a integridade cromossômica por associação a revestimentos protetores. Esses revestimentos seqüestram o prolongamento unifilamentar 3', que pode ter até 100 nucleotídeos de tamanho (Figura 7.27). Sem esse revestimento protetor, os filamentos bifilamentares dos cromossomos seriam confundidos com quebras bifilamentares pela célula e tratados como tais. Como você verá mais adiante, no Capítulo 15, as quebras bifilamentares são potencialmente muito perigosas porque podem resultar em instabilidade cromossômica que pode levar a câncer e a uma variedade de fenótipos associados ao envelhecimento. Por esse motivo, quando é detectada uma quebra bifilamentar, a célula responde de uma variedade de modos, dependendo, em parte, do tipo celular e da amplitude do dano. Por exemplo, a quebra bifilamentar pode ser fundida a outra quebra, ou a célula pode limitar o dano ao organismo parando a divisão celular (chamada senescência), ou iniciando uma via de morte celular (chamada apoptose).

Os telômeros são estruturas especializadas, nas pontas dos cromossomos, que contêm repetições em tandem de uma curta sequência de DNA que é adicionada à ponta 3' pela enzima telomerase. Os telômeros estabilizam os cromossomos evitando a perda de informação genômica após cada rodada de replicação do DNA, e associando-se a proteínas para formar um revestimento que "esconde" as pontas dos cromossomos da maquinaria de reparo de DNA da célula.

Ampliação de telômero

(a) AMPLIAÇÃO DO PROLONGAMENTO 3'



(b) REPLICAÇÃO DE FILAMENTO COMPLEMENTAR

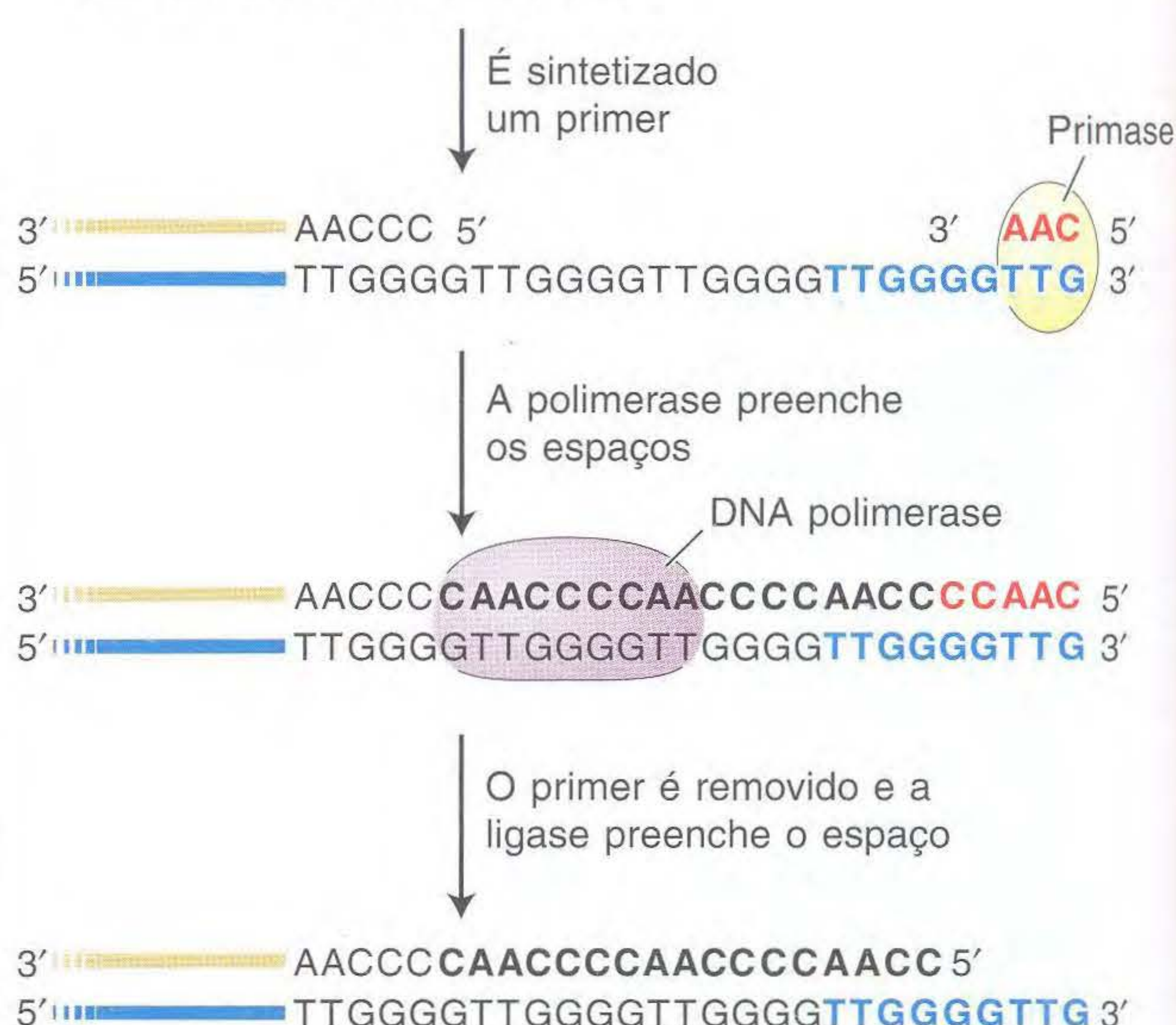


FIG. 7.26 A telomerase leva uma curta molécula de RNA (letras vermelhas) que atua como um molde para a adição de uma sequência complementar de DNA, que é adicionada ao prolongamento em 3' (letras azuis). Para adicionar outra repetição, a telomerase transloca-se para a ponta da repetição que foi adicionada. O prolongamento de 3' pode então servir como molde para a replicação convencional do DNA. [Lin Kah Wai. "Telomeres, Telomerase, and Tumorigenesis: A Review." *Medscape Gen Med*. 2004, 6(3):19. ©2004 Medscape.]

Estrutura do revestimento telomérico

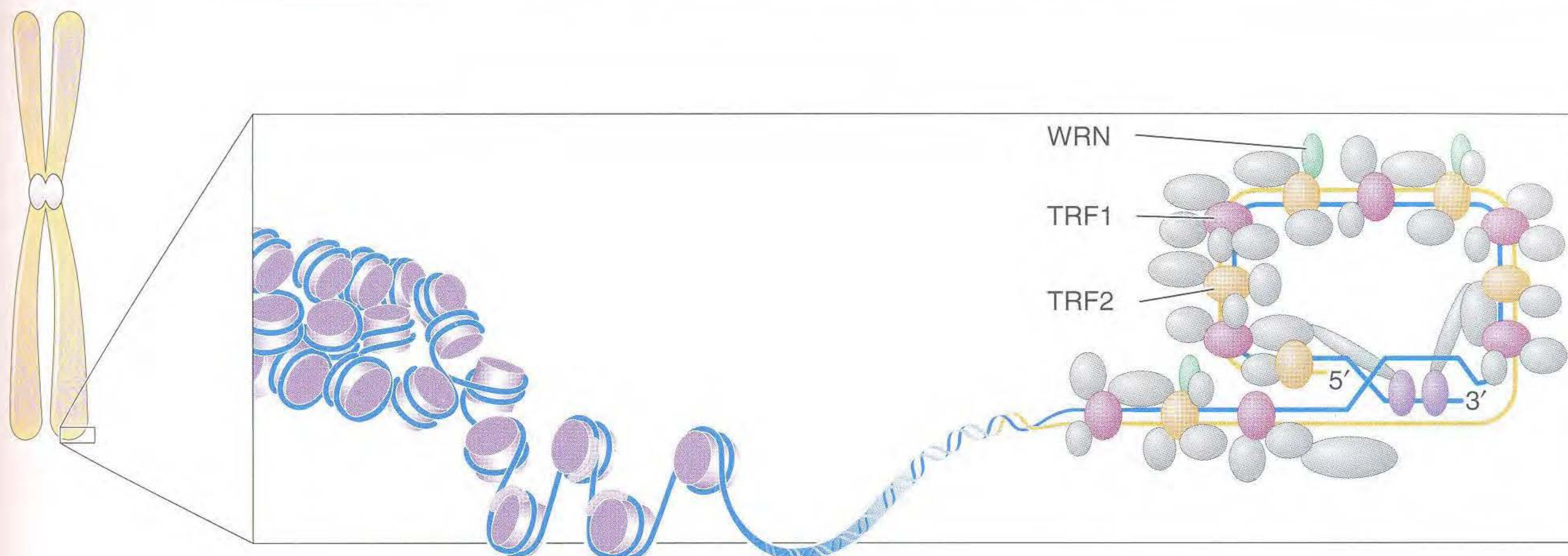


FIG. 7.27 Um “revestimento” protege o telômero na ponta de um cromossomo. O prolongamento 3' é “escondido” quando desloca um filamento de DNA em uma região onde as repetições teloméricas são bifilamentares. As proteínas TRF1 e TRF2 ligam-se a repetições teloméricas, e outras proteínas, incluindo WRN, ligam-se a TRF1 e TRF2, formando assim o revestimento telomérico protetor.

Surpreendentemente, embora a maioria das células germinativas tenham ampla telomerase, as células somáticas produzem muito pouca ou nenhuma telomerase. Por esse motivo, os cromossomos das células somáticas proliferativas ficam progressivamente mais curtos a cada divisão celular, até que a célula pára todas as divisões e entra em fase de senescência. Essa observação levou muitos pesquisadores a suspeitar que há uma ligação entre o encurtamento da polimerase e o envelhecimento. Os geneticistas, estudando doenças humanas que levam ao fenótipo de envelhecimento prematuro, recentemente descobriram evidências que apóiam tal associação. As pessoas com a síndrome de Werner sofrem o início precoce de muitos eventos relacionados ao envelhecimento, incluindo rugas na pele, cataratas, osteoporose, cabelos grisalhos e doença cardiovascular (Figura 7.28). Os estudos genéticos e bioquímicos descobriram que as pessoas afetadas têm telômeros mais curtos do que as pessoas normais devido a uma mutação em um gene chamado WRN; este codifica uma proteína (uma helicase) que é parte da estrutura de revestimento do telômero (veja Figura 7.27). Essa mutação supostamente perturba o telômero normal, resultando em instabilidade cromossômica e no fenótipo de envelhecimento prematuro. Os pacientes com outra síndrome de envelhecimento prematuro, chamada de disceratose congênita (DC), também têm telômeros mais curtos que as pessoas saudáveis da mesma idade, bem como possuem mutações nos genes necessários para a atividade da telomerase.

Os geneticistas também estão muito interessados nas associações entre telômeros e câncer. Ao contrário das células somáticas normais, a maioria das células cancero-

A síndrome de Werner provoca envelhecimento prematuro



FIG. 7.28 Uma mulher com síndrome de Werner aos 15 e 48 anos de idade. [International Registry of Werner Syndrome, www.wernersyndrome.org.]

sas têm atividade de telomerase. A habilidade em manter telômeros funcionais pode ser um dos motivos pelos quais as células cancerosas, mas não as normais, podem crescer em culturas de células por décadas, e são consideradas imortais. Como tal, muitas empresas farmacêuticas estão procurando focar essa diferença entre as células normais e as cancerosas desenvolvendo drogas que têm como alvo seletivo as células cancerosas inibindo a atividade da telomerase.

Resumo

O trabalho experimental sobre a natureza molecular do material hereditário demonstrou, conclusivamente, que o DNA (e não as proteínas, lipídios ou carboidratos) é de fato o material genético. Usando dados obtidos por outros, Watson e Crick

deduziram um modelo de dupla hélice com dois filamentos de DNA, enrolados um ao redor do outro, correndo com polaridade inversa. A ligação dos dois filamentos é baseada no pareamento entre adenina (A) e timina (T) e entre gua-

nina (G) e citosina (C). O primeiro par é mantido por duas pontes de hidrogênio; o segundo, por três.

O modelo de Watson-Crick mostra como o DNA pode se replicar de modo ordenado, um requisito principal do material genético. A replicação é feita semiconservativamente tanto em procariontes quanto em eucariontes. Uma dupla hélice é replicada para formar duas hélices idênticas, cada uma com seus nucleotídeos na ordem linear idêntica. Cada uma das duas novas duplas hélices é composta de um filamento polimerizado de DNA novo e um antigo.

A dupla hélice é desenrolada na forquilha de replicação, e os dois filamentos isolados assim produzidos servem como moldes para a polimerização de nucleotídeos livres. Os nucleotídeos são polimerizados pela enzima DNA polimerase, que adiciona nucleotídeos novos à ponta 3' de uma cadeia crescente de DNA. Como a adição é apenas nas pontas 3', a polimerização em um molde é contínua, produzindo o filamento contínuo, e no outro é descontínua em trechos curtos (fragmentos de Okazaki), produzindo o filamento chamado de lagging. A síntese do filamento contínuo (leading) e de cada fragmento de Okazaki é iniciada por um curto primer de RNA (sintetizado pela primase) que fornece uma ponta 3' para a adição de desoxirribonucleotídeos.

Os vários eventos que devem ocorrer com precisão e rapidez na forquilha de replicação são efetuados pelo replissomo,

uma máquina biológica. Esse complexo de proteína inclui duas unidades de DNA polimerase: uma para agir no filamento contínuo e outra para agir no filamento descontínuo. Desse modo, a síntese demora mais tempo e une os fragmentos de Okazaki em um filamento contínuo, e é temporariamente coordenada com a síntese menos complicada do filamento contínuo. Onde e quando ocorre a replicação são cuidadosamente controlados pela montagem ordenada do replissomo em alguns locais, ou origens, no cromossomo. Os genomas eucarióticos podem ter dezenas de milhares de origens. A montagem dos replissomos nessas origens pode ocorrer apenas em uma época específica do ciclo celular.

As pontas (telômeros) de cromossomos lineares apresentam um problema para o sistema de replicação porque há sempre um curto trecho em um filamento que não pode ter primer. A enzima telomerase adiciona várias seqüências curtas, repetitivas, para manter o tamanho. A telomerase leva um RNA curto que atua como o molde para a síntese das repetições teloméricas. As repetições teloméricas não-codificantes associam-se a proteínas para formar um revestimento telomérico. Os telômeros encurtam com a idade, nas células somáticas, porque a telomerase não é feita nessas células. As pessoas que têm telômeros defeituosos sofrem um envelhecimento prematuro.

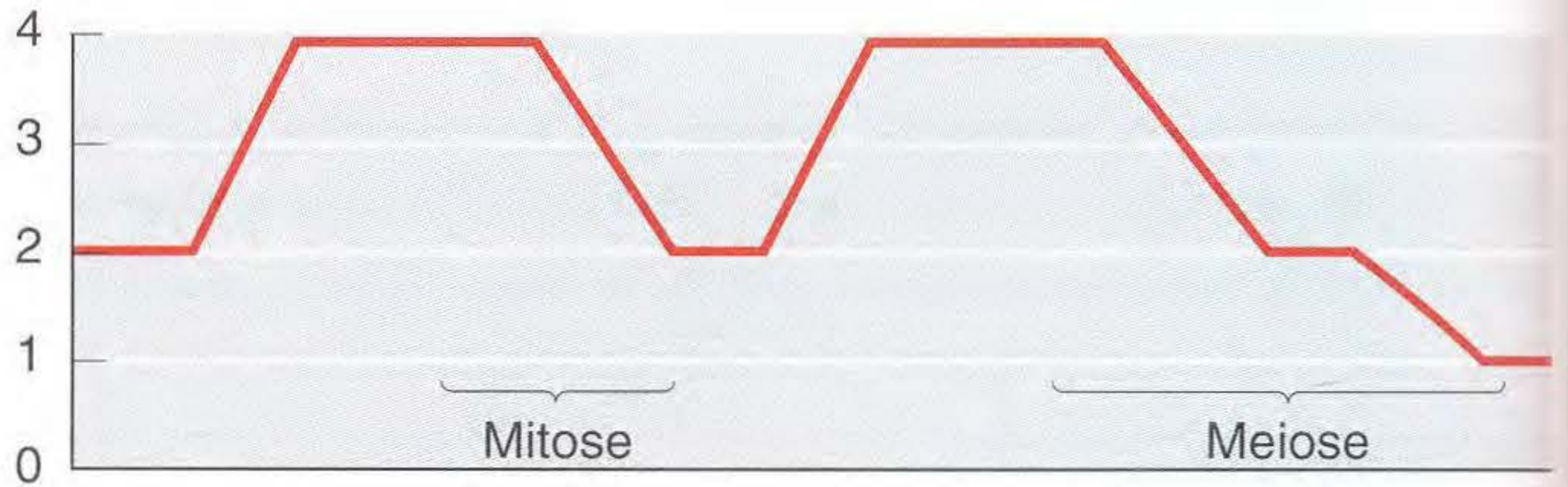
► Termos-chave

antígeno nuclear de proliferação celular (PCNA)	fosfato	primossomo
base	fragmento de Okazaki	proteína acessória
base complementar	grampo β	proteína de ligação unifilamentar (SSB)
código genético	helicase	purina
desoxirribose	holoenzima polimerase III (pol III)	replicação conservativa
DNA ligase	molde	replicação dispersiva
dupla hélice	molécula-filha	replicação semiconservativa
enzima distributiva	nucleossomo	replissomo
enzima processiva	nucleotídeo	sulco maior
estrutura teta (θ)	orientação em polaridade inversa	sulco menor
filamento contínuo (leading)	origem	telomerase
filamento descontínuo (lagging)	pirimidina	telômero
forquilha de replicação	primase	topoisomerase
	primer	

► Problemas Resolvidos

Problema resolvido 1. A mitose e a meiose foram apresentadas no Capítulo 2. Considerando o que foi mostrado neste capítulo quanto à replicação do DNA, desenhe um gráfico mostrando o conteúdo de DNA contra o tempo na célula que sofre mitose e, então, meiose. Suponha uma célula diplóide.

SOLUÇÃO



Problema resolvido 2. Se o conteúdo de GC de uma molécula de DNA é de 56%, quais são as percentagens das quatro bases (A, T, G e C) nessa molécula?

SOLUÇÃO

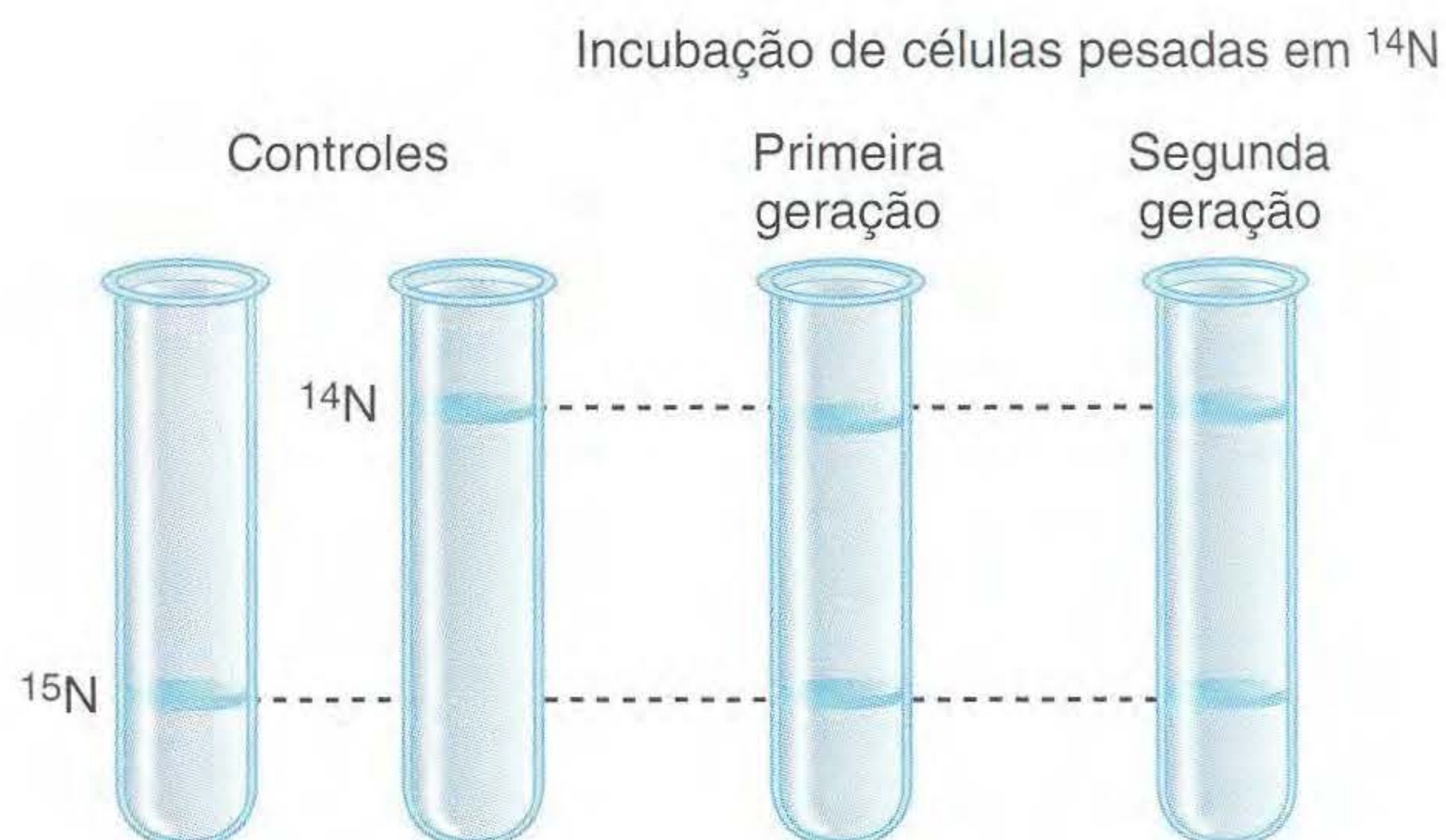
Se o conteúdo de GC é de 56%, então, como $G = C$, o conteúdo de G é de 28%, e o conteúdo de C é de 28%. O conteúdo de AT é $100 - 56 = 44\%$. Como $A = T$, o conteúdo de A é de 22%, e o conteúdo de T é de 22%.

Problema resolvido 3. Descreva o padrão esperado de bandas em um gradiente de CsCl para a replicação *conservativa* no experimento de Meselson-Stahl. Desenhe um diagrama.

SOLUÇÃO

Veja a Figura 7.13 para uma explicação adicional. Na replicação conservativa, se as bactérias são cultivadas na presença de ^{15}N e então mudadas para ^{14}N , uma molécula de DNA será totalmente ^{15}N após a primeira geração e a outra molécula

será toda ^{14}N , resultando em uma banda pesada e uma banda leve no gradiente. Após a segunda geração, o DNA ^{15}N irá liberar uma molécula toda com ^{15}N e uma molécula toda com ^{14}N , enquanto o DNA ^{14}N dará só DNA ^{14}N . Assim, serão gerados DNA só com ^{14}N ou DNA só com ^{15}N , novamente produzindo uma banda leve e uma banda pesada:



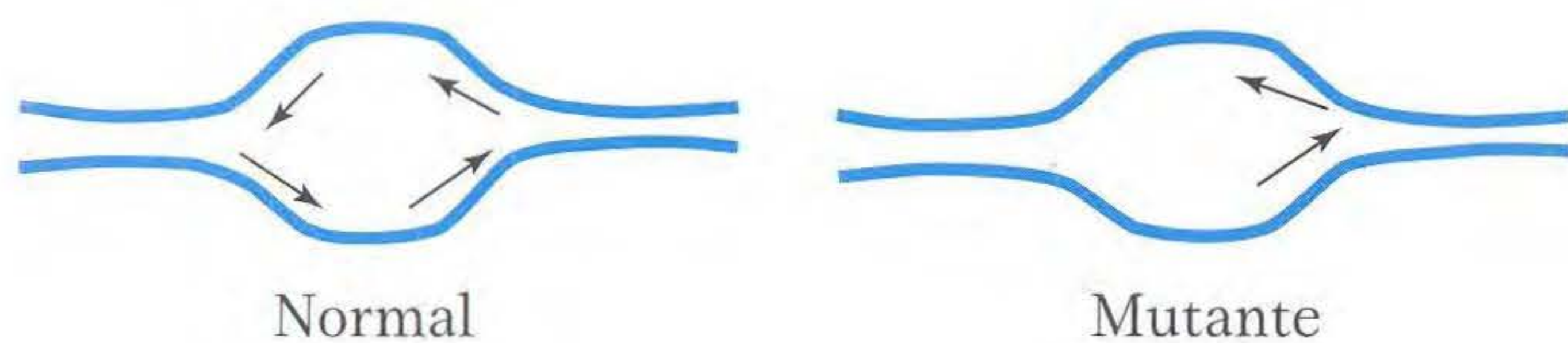
Problemas

PROBLEMAS BÁSICOS

1. Descreva os tipos de ligações químicas na dupla hélice de DNA.
2. Explique o que significam os termos *replicação conservativa* e *semiconservativa*.
3. Qual o significado de um *primer*, e por que os primers são necessários para a replicação do DNA?
4. O que são helicases e topoisomerases?
5. Por que a síntese de DNA é contínua em um filamento e descontínua no filamento oposto?
6. Se os quatro desoxinucleotídeos mostrassem pareamento inespecífico de bases (A com C, A com G, T com G, e assim por diante), a informação única contida em um gene seria mantida após várias rodadas de replicação? Explique.
7. Se as helicases estivessem faltando durante a replicação, o que aconteceria com o processo de replicação?
8. Ambos os filamentos de uma molécula de DNA são replicados simultaneamente de modo contínuo em um filamento e descontínuo no outro. Por que um filamento pode ser totalmente replicado (de ponta a ponta) antes que a replicação do outro tenha iniciado?
9. O que aconteceria se, no curso da replicação, as topoisomerases fossem incapazes de religar os fragmentos de DNA de cada filamento após a deslicoidização (relaxamento) da molécula de DNA?
10. Qual dos seguintes eventos ocorreria se a síntese do DNA fosse descontínua em ambos os filamentos?
 - a. Os fragmentos de DNA dos dois novos filamentos poderiam estar misturados, produzindo possíveis mutações.
 - b. A síntese de DNA não ocorreria, pois as enzimas apropriadas para fazer a replicação descontínua em ambos os filamentos não estariam presentes.
 - c. A síntese de DNA demoraria mais, porém, de outro modo, não haveria diferença notável.

d. A síntese de DNA não ocorreria, pois todo o tamanho do cromossomo teria que ser desenrolado antes que ambos os filamentos pudessem ser replicados de modo descontínuo.

11. Qual dos seguintes itens *não é* uma propriedade-chave do material hereditário?
 - a. Ele deve ser capaz de ser copiado com precisão.
 - b. Ele deve codificar a informação necessária para formar proteínas e estruturas complexas.
 - c. Ocasionalmente ele deve mutar.
 - d. Ele deve ser capaz de se adaptar a cada um dos tecidos do corpo.
12. É essencial que os primers de RNA nas pontas dos fragmentos de Okazaki sejam removidos e substituídos por DNA porque, de outro modo, qual dos seguintes eventos resultaria?
 - a. O RNA pode não ser precisamente lido durante a transcrição, interferindo assim na síntese de proteínas.
 - b. O RNA mais provavelmente conteria erros porque a primase não tem função de revisão.
 - c. Os trechos de RNA desestabilizariam e começariam a se fragmentar em ribonucleotídeos, criando assim espaços na sequência.
 - d. Os primers de RNA provavelmente formariam pontes de hidrogênio uns com os outros, formando estruturas complexas que podem interferir na formação apropriada da hélice de DNA.
13. As polimerases geralmente adicionam apenas cerca de 10 nucleotídeos a um filamento de DNA antes de se dissociarem. Entretanto, durante a replicação, a DNA pol III pode adicionar dezenas de milhares de nucleotídeos a uma forquilha em movimento. Como é feita essa adição?
14. Em cada origem de replicação, existem duas forquilhas bidirecionais de replicação. Qual dos seguintes itens aconteceria se um mutante surgisse tendo apenas uma forquilha funcional por bolha de replicação? (Veja diagrama.)



Normal

Mutante

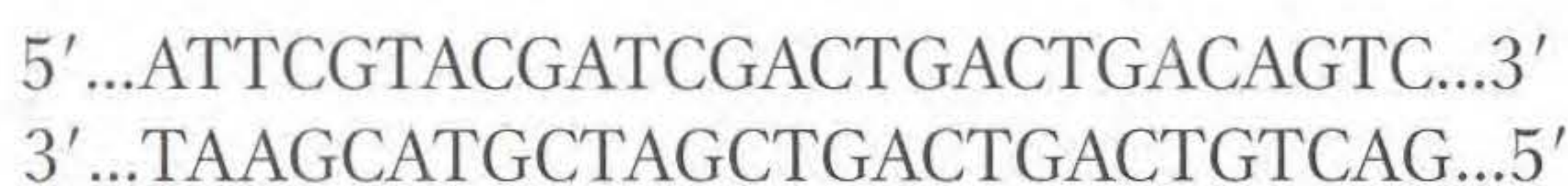
- a. Nenhuma mudança na replicação.
 b. A replicação ocorreria apenas em uma metade do cromossomo.
 c. A replicação seria completa apenas no filamento contínuo.
 d. A replicação teria o dobro do tamanho.
15. Em uma célula diplóide na qual $2n = 14$, quantos telômeros existem em cada uma das seguintes fases do ciclo celular: (a) G1; (b) G2; (c) prófase mitótica; (d) telófase mitótica?
16. Se a timina constitui 15% das bases em uma molécula específica de DNA, que percentagem das bases é de citosina?
17. Se o conteúdo GC de uma molécula de DNA é 48%, quais as percentagens das quatro bases (A, T, G e C) nessa molécula?
18. Suponha que determinado cromossomo bacteriano tem uma origem de replicação. Sob algumas condições de divisão celular rápida, a replicação pode começar a partir da origem antes que o ciclo de replicação precedente esteja completo. Quantas forquilha de replicação estariam presentes nessas condições?
19. Uma molécula com a composição



é replicada em uma solução de trifosfatos do nucleosídeo adenina com todos os seus átomos de fósforo na forma do isótopo radioativo ^{32}P . Ambas as moléculas-filhas serão radioativas? Explique. Então repita a pergunta para a molécula



20. O experimento de Meselson e Stahl funcionaria se tivessem sido usadas células eucarióticas diplóides?
21. Considere o seguinte segmento de DNA, que é parte de uma molécula muito maior constituindo um cromossomo:

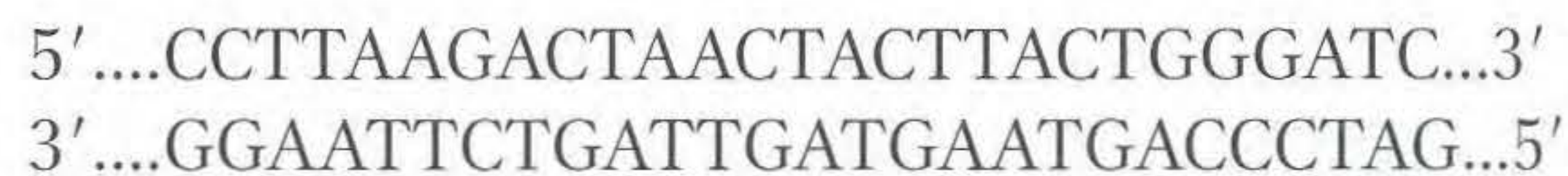


Se a DNA polimerase começar a replicar esse segmento a partir da direita,

- a. Qual será o molde para o filamento contínuo?

- b. Desenhe a molécula quando a DNA polimerase está na metade do caminho desse segmento.
 c. Desenhe as duas moléculas-filhas completas.
 d. O seu diagrama da parte b é compatível com a replicação bidirecional de uma única origem, o modo usual de replicação?

22. As DNA polimerases são posicionadas sobre o seguinte segmento de DNA (que é parte de uma molécula maior) e movem-se da direita para a esquerda. Se for suposto que um fragmento de Okazaki é feito desse segmento, qual será a sequência do fragmento? Marque suas pontas 5' e 3'.



23. Os cromossomos de *E. coli* nos quais cada átomo de nitrogênio é marcado (isto é, cada átomo de nitrogênio é o isótopo pesado ^{15}N em vez do isótopo normal ^{14}N) são postos para se replicar em um ambiente no qual todo o nitrogênio é ^{14}N . Usando uma linha contínua para representar uma cadeia polinucleotídica pesada e uma linha tracejada para uma cadeia leve, esquematize cada uma das seguintes descrições:

- a. O cromossomo parental pesado e os produtos da primeira replicação após a transferência para um meio com ^{14}N , supondo que o cromossomo é uma dupla hélice de DNA e que a replicação é semiconservativa.
 b. Repita a parte a, mas agora suponha que a replicação é conservativa.
 c. Repita a parte a, mas suponha que o cromossomo é de fato duas duplas hélices lado a lado, cada uma das quais se replica semiconservativamente.
 d. Repita a parte c, mas suponha que cada dupla hélice lado a lado replica-se conservativamente e que a replicação geral do cromossomo é semiconservativa.
 e. Se os cromossomos-filhos da primeira divisão em ^{14}N são centrifugados em um gradiente de densidade de cloreto de cério e é obtida uma única banda, quais as possibilidades de as partes a a d serem excluídas? Reconsidere o experimento de Meselson e Stahl: O que ele prova?

24. Uma estudante do laboratório de Griffiths descobriu três amostras de células marcadas "A", "B" e "C". Sem saber o que cada amostra continha, a estudante decidiu tentar injetar essas amostras em alguns de seus camundongos, tanto isoladamente quanto em combinações, para ver se podia determinar o que cada amostra continha. Ela observou as respostas dos camundongos infectados após um período de incubação e colheu amostras de sangue de cada grupo de teste para procurar a possível presença de células infectantes. Ela registrou suas observações no quadro seguinte. Com a suposição de que cada amostra continha apenas uma coisa em forma pura, o que você acha que havia nas amostras "A", "B" e "C"?

Amostra injetada	Resposta do camundongo	Tipo de células recuperadas de camundongos
A	morte	células S vivas
B	nenhuma	nenhuma
C	nenhuma	células R vivas
A + B	morte	células S vivas
A + C	morte	células R e S vivas
B + C	morte	células S vivas
A + B + C	morte	células S vivas

25. Se, no experimento do Problema 24, as proteínas do revestimento celular foram os fatores transformantes, que resultados você esperaria? Complete o quadro seguinte. (Lembre que as amostras são as mesmas que no Problema 24.)

Amostra injetada	Resposta de camundongos	Tipo de células recuperadas de camundongos
A		
B		
C		
A + B		
A + C		
B + C		
A + B + C		

PROBLEMAS DESAFIADORES

26. Se uma mutação que inativa a telomerase ocorre em uma célula (atividade de telomerase na célula = zero), que resultado você espera?
27. No planeta Rama, o DNA é de seis tipos de nucleotídeos: A, B, C, D, E e F. Os tipos A e B são chamados *marzines*, C e D são *orsines* e E e F são *pirines*. As seguintes regras são válidas em todos os DNA de Rama:

Total de marzines = total de orsines = total de pirines

A = C = E
B = D = F

- a. Prepare um modelo para a estrutura do DNA de Rama.
- b. Em Rama, a mitose produz três células-filhas. Tendo em mente esse fato, proponha um padrão de replicação para seu modelo de DNA.
- c. Considere o processo de meiose em Rama. Que comentários ou conclusões você pode sugerir?
28. Se você extrai o DNA do colifago ϕ X174, você verá que sua composição é 25% A, 33% T, 24% G e 18% C. Essa composição faz sentido com relação às regras de Chargaff? Como você interpretaria esse resultado? Como tal fago replica seu DNA?

RNA: Transcrição e Processamento

8



A RNA polimerase em ação. Uma RNA polimerase bem pequena (azul), feita pelo bacteriófago T7, transcreve o DNA em um filamento de RNA (vermelho). A enzima separa a dupla hélice de DNA (amarelo, laranja) expondo o filamento-molde a ser copiado em RNA. [David S. Goodsell, Scripps Research Institute.]

Perguntas Fundamentais

- Como a estrutura do RNA difere da do DNA?
- Quais são as diferentes classes de RNA em uma célula?
- Como a RNA polimerase é corretamente posicionada para começar a transcrição?
- Como o RNA eucariótico é modificado antes de deixar o núcleo?
- Por que a descoberta dos íntrons auto-removíveis é considerada tão importante?

Tópicos

- 8.1 RNA
- 8.2 Transcrição
- 8.3 Transcrição em eucariontes
- 8.4 RNA funcionais

Usando seus conhecimentos recém-adquiridos das seqüências de DNA de genomas inteiros, os cientistas foram capazes de determinar o número de genes em vários organismos, tanto simples quanto complexos. No início não houve surpresas; a bactéria *Escherichia coli* tem cerca de 3.200 genes, o eucarionte unicelular, a levedura *Saccharomyces cerevisiae*, tem cerca de 6.300 genes e a mosca das frutas *Drosophila melanogaster* tem cerca de 13.600 genes. O aumento de complexidade foi suposto como requerendo mais genes, e, assim, as primeiras estimativas foram de que nosso genoma tivesse 100.000 genes. Em uma conferência enfocada nas pesquisas genômicas em 2000, os cientistas começaram com uma reunião informal, chamada GeneSweep, que seria vencida pela pessoa que obtivesse a previsão mais próxima do número real de genes no genoma humano. Os números variaram de cerca de 26.000 a cerca de 150.000 genes.

Com a liberação da primeira sequência rascunho, foi anunciado um vencedor. Surpreendentemente, o que entrou com a mais baixa estimativa, 25.947 genes, foi declarado o vencedor. Como pode o *Homo sapiens*, com seu cérebro complexo e sofisticado sistema imune, ter apenas o dobro dos genes que um verme nematelminto e o mesmo número de genes que o primeiro genoma de planta sequenciado, a praga *Arabidopsis thaliana*? A resposta a essa pergunta tem a ver com a marcante descoberta feita no final dos anos de 1970. Nessa época, as proteínas de organismos superiores foram descobertas como sendo codificadas no DNA não como trechos contínuos (como são nas bactérias), mas em pedaços. Assim, os genes de eucariontes superiores são geralmente compostos de pedaços chamados **éxons** (para regiões expressas), que codificam partes das proteínas, e pedaços chamados **íntrons** (para regiões intercalares), que separam os éxons. Como você verá neste capítulo, uma cópia de RNA contendo tanto éxons quanto íntrons é sintetizada de um gene. Uma máquina biológica (chamada **spliceossomo**) remove os íntrons e une os éxons (em um processo chamado de **recomposição do RNA**) para produzir um RNA final que contém a informação contínua necessária para sintetizar uma proteína.

O que os éxons e os íntrons têm a ver com a baixa contagem de genes humanos? Por enquanto, é suficiente dizer que o RNA transcrito de um gene pode ser recomposto de modos alternativos. Embora tenhamos apenas cerca de 25.000 genes, esses genes codificam mais de cerca de 100.000 proteínas, graças ao processo de **recomposição alternativa** do RNA.

Neste capítulo, vemos as primeiras etapas na transferência de informações dos genes para os produtos gênicos. Dentro da sequência do DNA do genoma de qualquer organismo está codificada a informação que especifica cada um dos produtos gênicos que o organismo pode fazer. Essas sequências de DNA também contêm informações que especificam quando, onde e como grande parte do produto é feita. Entretanto, essa informação é estática, inserida na sequência do DNA. Para usar a informação, deve ser sintetizada uma molécula intermediária que é uma cópia de um gene distinto com o uso da sequência de DNA como um guia. Essa molécula é o RNA, e o processo de sua síntese a partir do DNA é chamado de **transcrição**.

A transferência de informação de um gene para o produto gênico ocorre em várias etapas. A primeira etapa, que é o foco deste capítulo, é copiar (*transcrever*) a informação em um filamento de RNA com o uso do DNA como um guia, ou molde. Em procariontes, a informação no RNA é quase imediatamente convertida em uma cadeia de aminoácidos (polipeptídica) por um processo chamado de **tradução**. Essa segunda etapa é o foco do Capítulo 9. Nos eucariontes, a transcrição e a tradução são separadas espacialmente: a transcrição ocorre no núcleo e a tradução no citoplasma. Entretanto, antes que os RNA estejam prontos para serem transportados para o citoplasma para tradução, eles sofrem um amplo processamento, incluindo a remoção dos íntrons e a adição de um revestimento (*cap*) especial em 5' e uma cauda em 3' de nucleotídeos adenina. Um RNA totalmente processado é chamado de **RNA mensageiro** (mRNA).

A função do DNA e do RNA é baseada em dois princípios:

1. A complementariedade de bases é responsável por determinar a sequência de um novo filamento de DNA na replicação e do RNA transcrito na transcrição. Pelo ajuste da complementariedade de bases, o DNA é replicado e a informação codificada no DNA passa para o RNA (e, finalmente, para a proteína).
2. Algumas proteínas reconhecem determinadas sequências de bases no DNA. Essas proteínas de ligação a ácidos nucleicos unem-se a essas sequências e atuam nelas.

Veremos esses dois princípios em funcionamento pelas discussões detalhadas da transcrição e tradução que seguem neste capítulo e nos capítulos a vir.

As transações de DNA e RNA ocorrem pelo pareamento de bases complementares e a ligação de várias proteínas a sítios específicos no DNA ou RNA.

► 8.1 RNA

Os primeiros pesquisadores tiveram bons motivos para pensar que a informação não é transferida diretamente do DNA para a proteína. Em uma célula eucariótica, o DNA é encontrado no núcleo, enquanto as proteínas são sintetizadas no citoplasma. É necessário um intermediário.

Os primeiros experimentos sugerem um RNA intermediário

Em 1957, Elliot Volkin e Lawrence Astrachan fizeram uma significativa observação. Eles descobriram que uma das mais marcantes mudanças moleculares que ocorrem quando *E. coli* é infectada pelo fago T2 é um rápido surto de síntese de RNA. Além disso, esse RNA induzido por fago “é renovado” rapidamente; isto é, seu tempo de vida é curto, da ordem de minutos. Seu surgimento e desaparecimento rápido sugerem que o RNA pode ter algum papel na expressão do genoma de T2 necessário para fazer mais partículas de vírus.

Volkin e Astrachan demonstraram a renovação rápida do RNA usando um protocolo chamado de **experimento de pulso-caça**. Para fazer um experimento de pulso-caça, as bactérias infectadas são primeiro alimentadas (pulso) com uracil radioativo (uma molécula necessária para a síntese de RNA, mas não de DNA). Qualquer RNA sintetizado nas bactérias a partir de então é “marcado” com uracil radioativo prontamente detectável. Após um curto período de incubação, uracil radioativo é removido e substituído (caça) por uracil não-radioativo. Esse procedimento “caça” a marcação para fora do RNA, pois, à medida que o RNA se degrada, apenas os precursores não-marcados estão disponíveis para sintetizar novas moléculas de RNA (os nucleotídeos marcados são “diluídos” pelo grande excesso de uracil não-marcado adicionado na caça). O RNA recuperado logo após o pulso é marcado, mas o recuperado um pouco mais tarde após a

caça não é marcado, indicando que o RNA tem um tempo de vida muito curto.

Um experimento similar pode ser feito com células eucarióticas. As células são primeiro “pulsadas” com uracil radioativo e, após um curto período de tempo, são transferidas para um meio com uracil não-marcado. Em amostras colhidas após o pulso, a maioria da marcação está no núcleo. Em amostras colhidas após a caça, o RNA marcado é encontrado no citoplasma (Figura 8.1). Aparentemente, em eucariontes, o RNA é sintetizado no núcleo e, então, move-se para o citoplasma, onde são sintetizadas as proteínas. Assim, o RNA é um bom candidato para um intermediário de transferência de informação entre o DNA e a proteína.

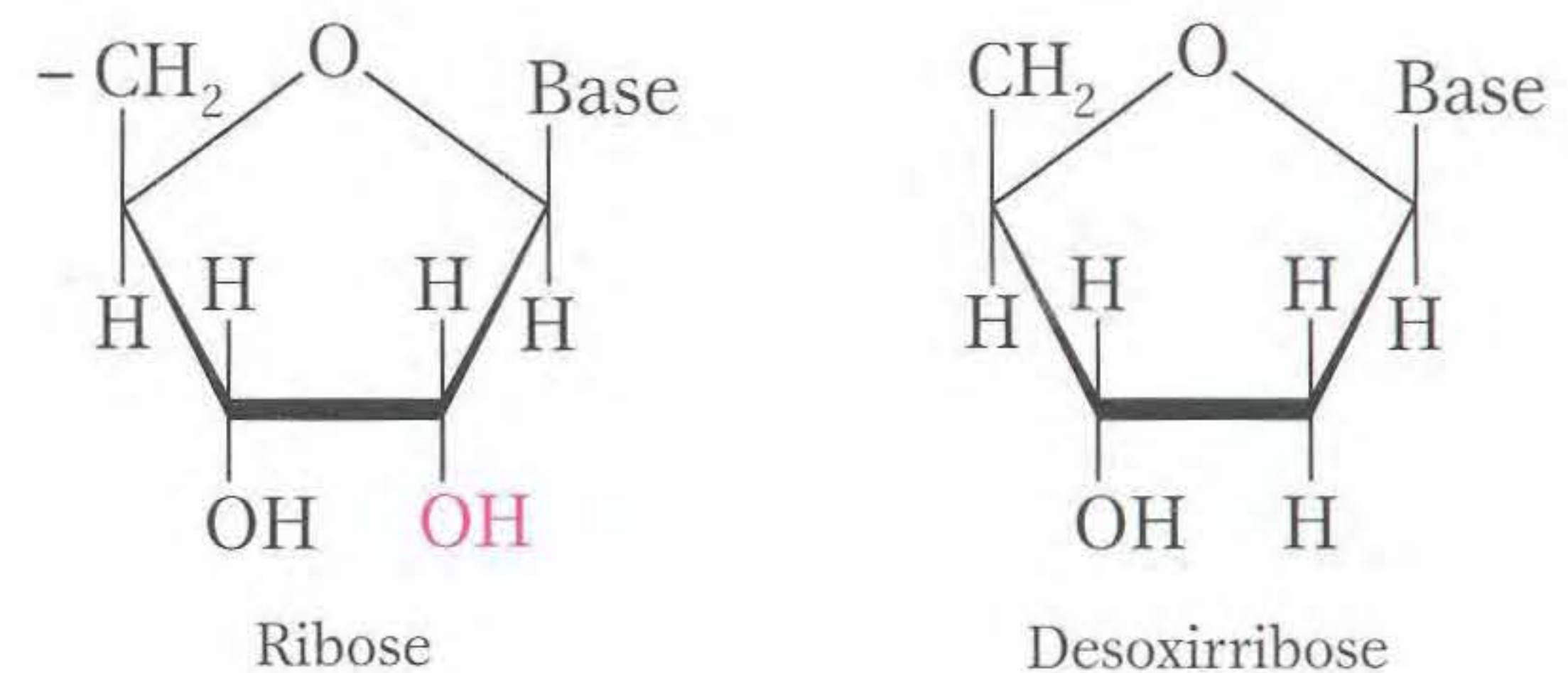
Propriedades do RNA

Consideremos as características gerais do RNA. Embora tanto o RNA quanto o DNA sejam ácidos nucleicos, o RNA difere do DNA de vários modos importantes:

1. O RNA é geralmente uma cadeia de nucleotídeos unifilamentares, não uma dupla hélice como o DNA. Uma consequência é que o RNA é mais flexível e pode formar uma variedade muito maior de formas moleculares tridimensionais complexas do que pode o DNA bifilamentar. Um filamento de RNA pode se dobrar de tal modo que algu-

mas de suas próprias bases podem fazer par com outras. Tal pareamento *intramolecular* de bases é um determinante importante da forma do RNA.

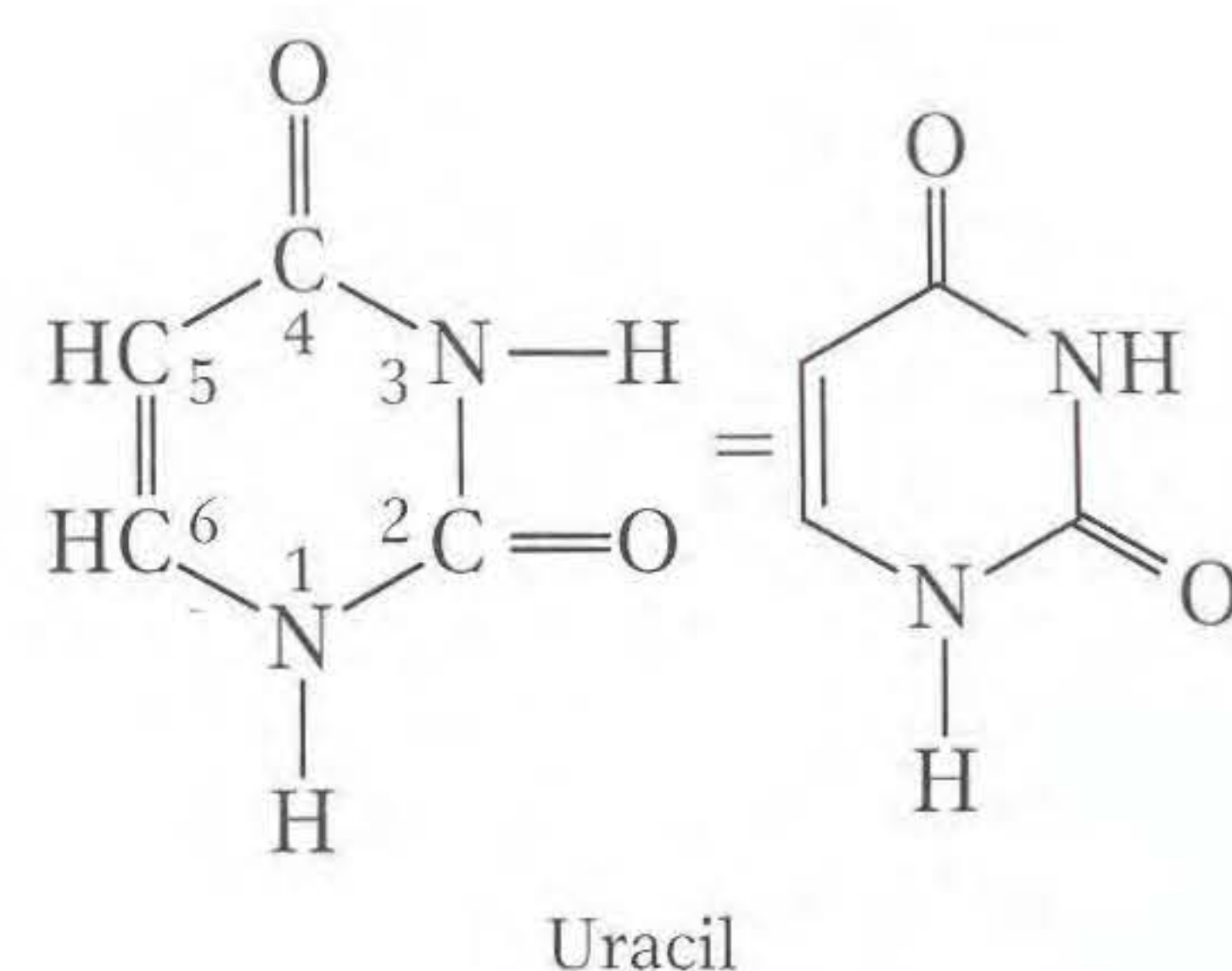
2. O RNA tem o açúcar **ribose** em seus nucleotídeos, em vez da desoxirribose encontrada no DNA. Como os nomes sugerem, os dois açúcares diferem na presença ou ausência apenas de um átomo de oxigênio. O açúcar do RNA contém um grupo hidroxila (OH) ligado ao átomo de carbono 2', enquanto o açúcar DNA tem apenas um átomo de hidrogênio ligado ao átomo de carbono 2'.



Como você verá neste capítulo, a presença do grupo hidroxila no átomo de carbono 2' facilita a ação do RNA em muitos processos celulares importantes.

Com um filamento individual de DNA, um filamento de RNA é formado de um arcabouço de fosfato e açúcar, com uma base covalentemente ligada à posição 1' em cada ribose. As ligações açúcar-fosfato são feitas nas posições 5' e 3' do açúcar, como no DNA; logo, uma cadeia de RNA terá uma ponta 5' e uma 3'.

3. Os nucleotídeos de RNA (chamados de ribonucleotídeos) contêm as bases adenina, guanina e citosina, mas a base pirimidínica **uracil** (abreviada por **U**) está presente no lugar de timina.



Uracil forma pontes de hidrogênio com adenina do mesmo modo que a timina faz. A Figura 8.2 mostra os quatro nucleotídeos encontrados no RNA.

Além disso, uracil é capaz de fazer pareamento de bases com G. As bases U e G formam pares de bases apenas durante o dobramento do RNA e não durante a transcrição. As duas pontes de hidrogênio que podem ser formadas entre U e G são mais fracas que as duas que se formam entre U e A. A habilidade de U parear com A e G é o principal motivo por que o RNA pode formar estruturas complicadas, muitas das quais são importantes em processos biológicos.

4. O RNA, como as proteínas, mas não como o DNA, pode catalisar reações biológicas. O nome **ribozima** foi criado para as moléculas de RNA que funcionam como enzimas proteicas.

O mRNA eucariótico move-se do núcleo para o citoplasma

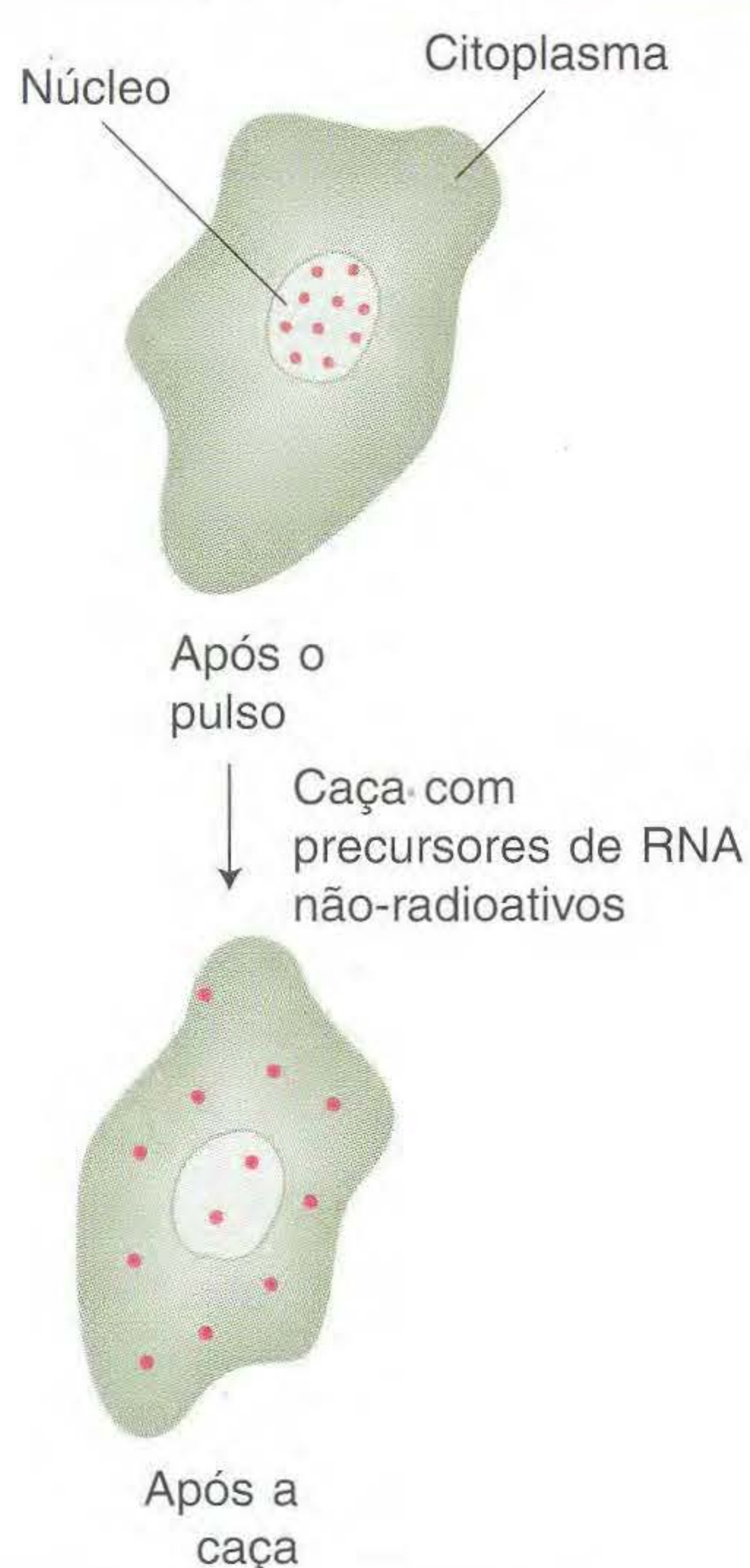
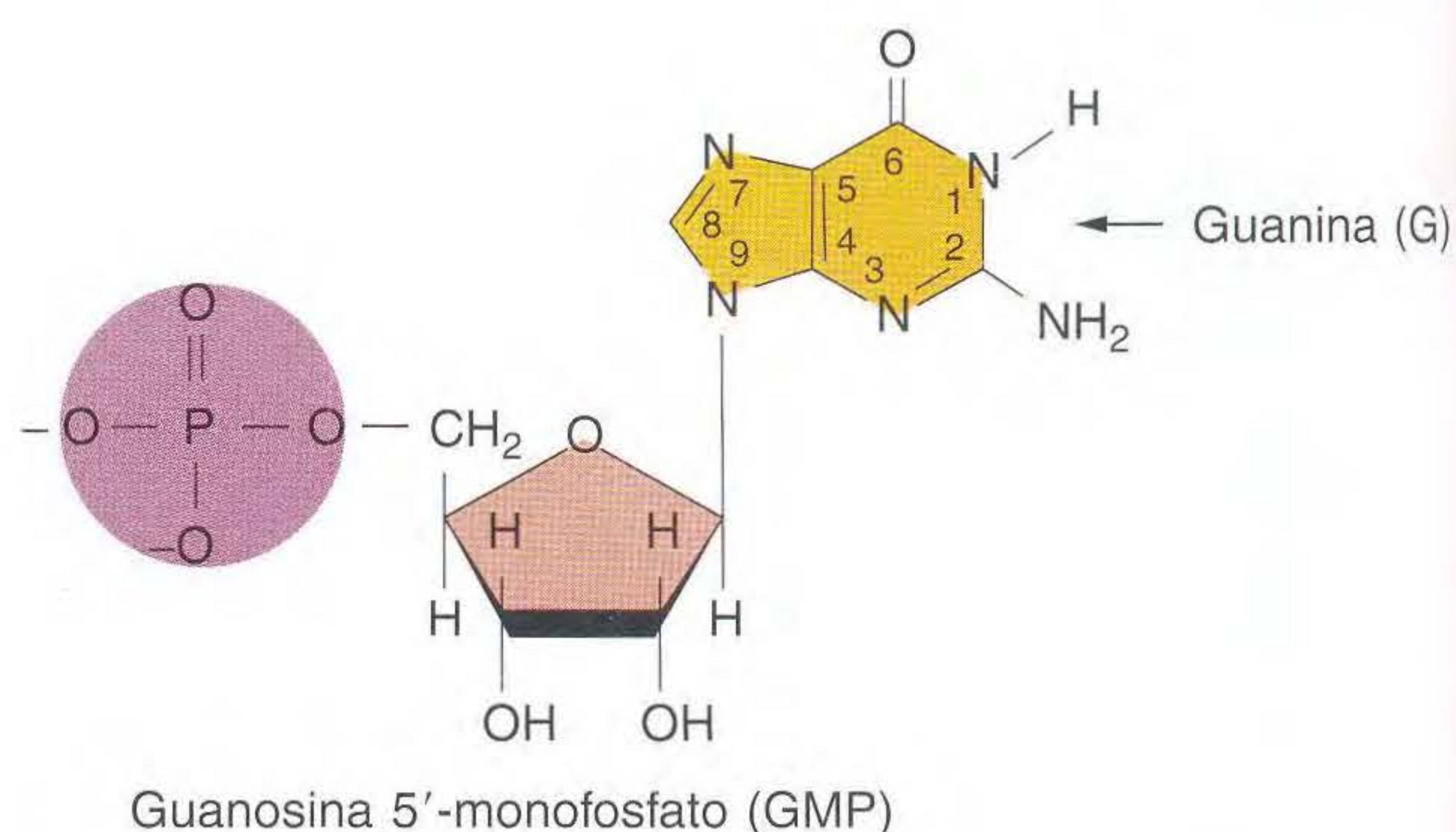
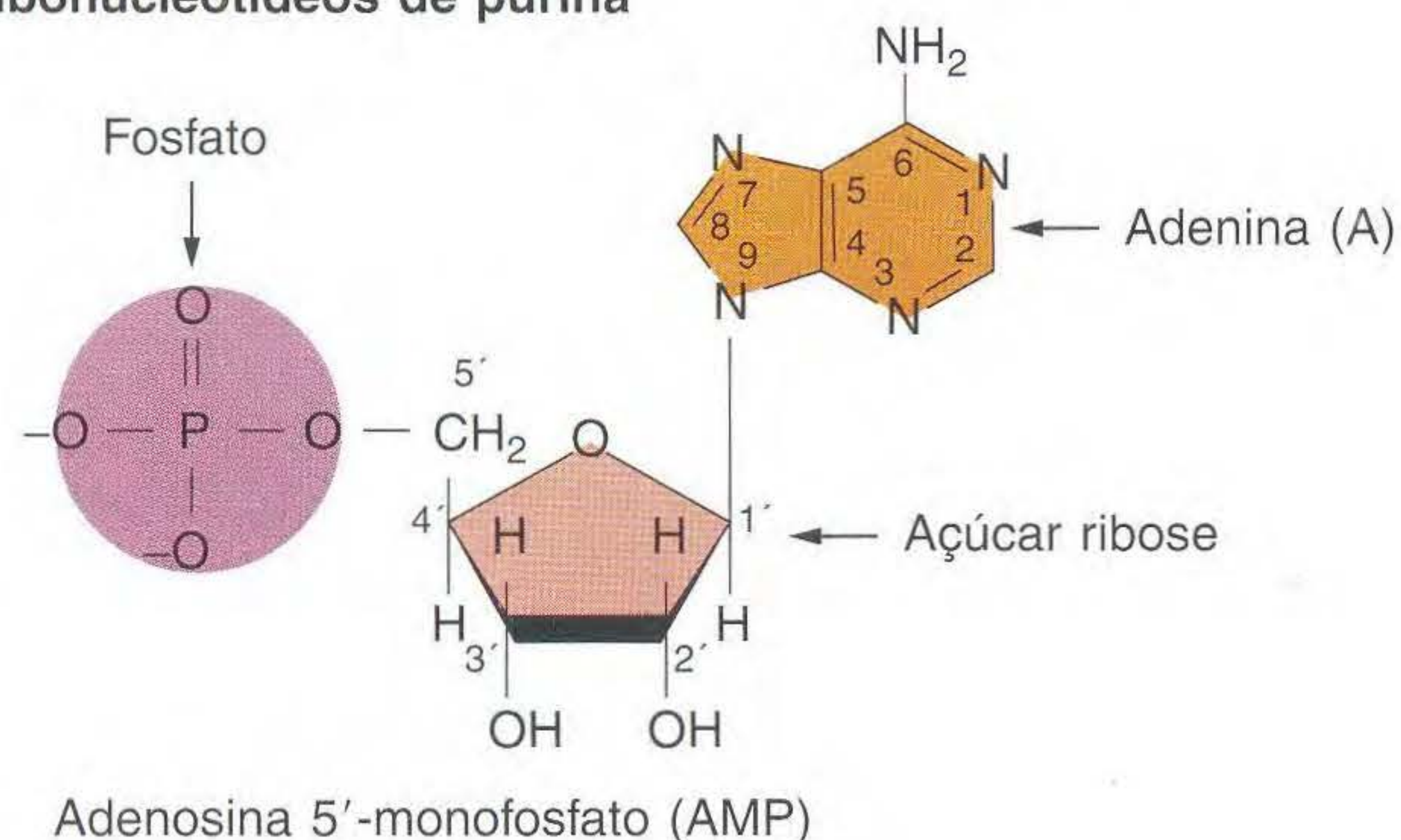


FIG. 8.1 O experimento de pulso-caça mostrou que o mRNA move-se para o citoplasma. As células são rapidamente cultivadas em uracil radioativo para marcar o RNA recém-sintetizado (pulso). As células são lavadas para remover uracil radioativo e, então, são cultivadas em um excesso de uracil não-radioativo (caça). Os pontos vermelhos indicam a localização do RNA contendo uracil radioativo com o tempo.

Os quatro ribonucleotídeos encontrados no RNA

Ribonucleotídeos de purina



Ribonucleotídeos de pirimidina

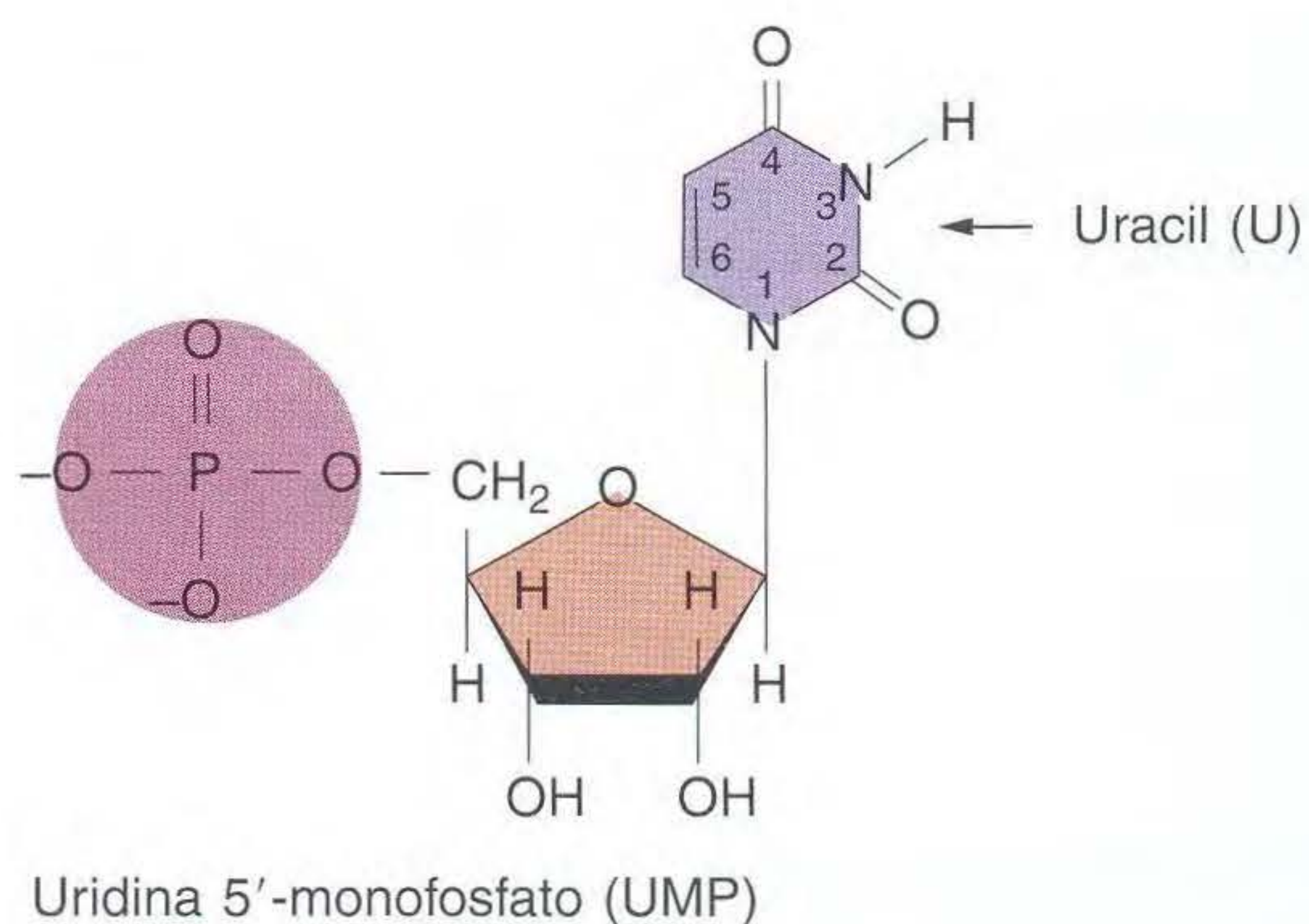
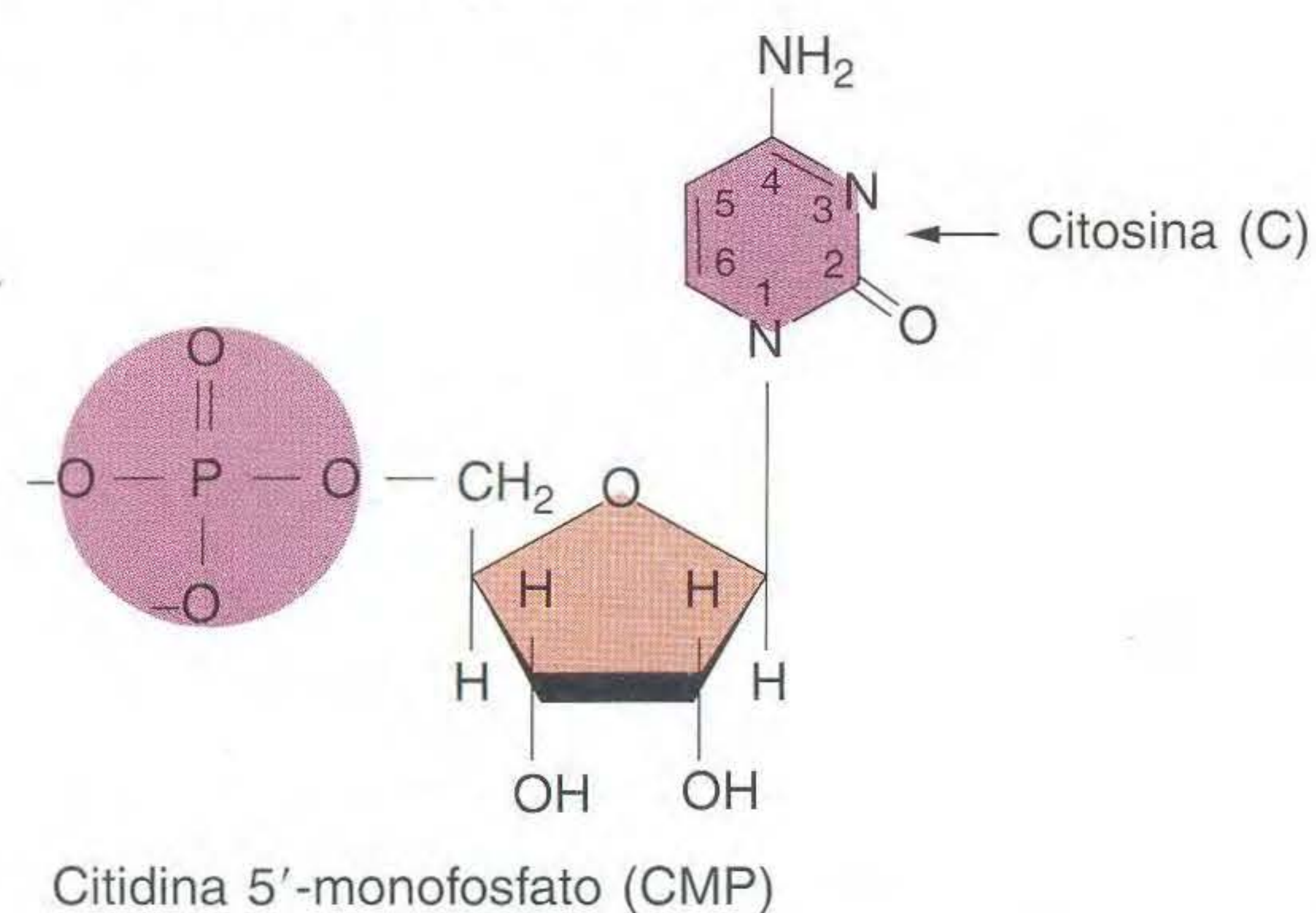


FIG. 8.2

Classes de RNA

Os RNA podem ser agrupados em duas classes gerais. Uma classe de RNA codifica a informação necessária para fazer cadeias polipeptídicas (proteínas). Chamamos essa classe de **RNA mensageiro (mRNA)** porque, como um mensageiro, esses RNA servem como intermediários e passam a informação do DNA para a proteína. Chamamos as outras classes de **RNA funcionais** porque o RNA não codifica a informação para fazer proteínas. Em vez disso, o próprio RNA é o produto funcional final. Como o RNA é uma molécula tão versátil, ele pode atender a muitos papéis funcionais e participar em uma variedade de processos celulares.

RNA mensageiro As etapas pelas quais um gene influencia o fenótipo são chamadas de *expressão gênica*. Para a grande maioria dos genes, o RNA transcrito é apenas um intermediário necessário para a síntese de uma proteína, que é o produto final funcional que influencia o fenótipo.

RNA funcional À medida que sabemos mais sobre os detalhes íntimos da expressão gênica e regulação, torna-se aparente que os RNA funcionais caem em uma variedade de classes e têm papéis diversos. Novamente, é importante enfatizar que os RNA funcionais são ativos como RNA; eles nunca são traduzidos em polipeptídeos.

As principais classes de RNA funcionais contribuem para várias etapas na transferência da informação do DNA para a proteína, no processamento de outros RNA e na regulação do RNA e níveis de proteínas na célula. Duas dessas classes de RNA funcionais são encontradas tanto em procariontes quanto eucariontes: RNA transportadores e RNA ribossômicos.

- **RNA transportador (tRNA)** são moléculas responsáveis por levar o aminoácido correto para o mRNA no processo de tradução.
- **RNA ribossômico (rRNA)** são moléculas dos principais componentes dos ribossomos, que são grandes máquinas macromoleculares que guiam a montagem da cadeia de aminoácidos pelo mRNA e tRNA.

Toda a coleção de tRNA e rRNA é codificada por um pequeno número de genes (algumas dezenas a centenas, no máximo). Entretanto, embora os genes que os codificam estejam em pequeno número, os rRNA respondem por uma grande percentagem do RNA na célula, porque eles são tanto estáveis quanto transcritos em muitas cópias.

Outra classe de RNA funcionais participam do processamento do RNA e são específicos dos eucariontes:

- **Pequenos RNA nucleares (snRNA)** são parte de um sistema que processa os RNA transcritos nas células eucarióticas. Alguns snRNA unem-se a várias subunidades

proteicas para formar o complexo ribonucleoproteico de processamento (o *spliceossomo*) que remove íntrons dos mRNA eucarióticos.

Finalmente, uma grande fração dos genomas eucarióticos é transcrita em um grupo diverso de RNA funcionais que participam na regulação da expressão gênica em muitos níveis. Duas classes desse RNA funcionais, microRNA e pequenos RNA de interferência, podem ser codificadas por grandes partes dos genomas eucarióticos:

- **MicroRNA (miRNA)** recentemente foram reconhecidos pelos cientistas como tendo um amplo papel na regulação da quantidade de proteínas produzidas por muitos genes eucarióticos.
- **Pequenos RNA de interferência (siRNA)** ajudam a proteger a integridade dos genomas de plantas e animais. Os pequenos RNA de interferência inibem a produção de vírus e impedem a dispersão de elementos de transposição para outros loci cromossômicos.

Existem duas classes gerais de RNA: as que codificam proteínas (mRNA) e as que são funcionais como RNA. Os RNA funcionais participam de uma variedade de processos celulares, incluindo a síntese de proteínas (tRNA, rRNA), RNA de processamento (snRNA), regulação da expressão gênica (miRNA) e defesa do genoma (siRNA).

8.2 Transcrição

A primeira etapa na transferência da informação do gene para a proteína é produzir um filamento de RNA cuja sequência de bases é complementar à sequência de bases de um segmento de DNA, às vezes seguido de modificação desse RNA para prepará-lo para seus papéis celulares específicos. Assim, o RNA é produzido por um processo que copia a sequência de nucleotídeos do DNA. Como esse processo é remanescente da transcrição (cópia) de palavras escritas, a síntese do RNA é chamada de **transcrição**. O DNA é dito sendo transcrito em RNA, e o RNA é chamado de **transcrito**.

Visão geral: o DNA como molde para a transcrição

Como a informação codificada na molécula de DNA é transferida para o RNA transcrito? A transcrição é baseada no pareamento complementar de bases. Considere a transcrição de um segmento cromossômico que constitui um gene. Primeiro, os dois filamentos da dupla hélice de DNA separam-se localmente, e um dos filamentos separado atua como um **molde** para a síntese de RNA. No cromossomo em geral, ambos os filamentos de DNA são usados como moldes; *mas, em qualquer gene, apenas um filamento é usado* e, nesse gene, é sempre o mesmo filamento (Figura 8.3). Em seguida, os ribonucleotídeos que foram quimicamente sintetizados em outra parte da célula formam pares estáveis com suas bases complementares no molde. O ribonucleotídeo A faz par com T no DNA, G com C, C com G e U com A. Cada ribonucleo-

Filamentos opostos de DNA podem servir como moldes para o RNA

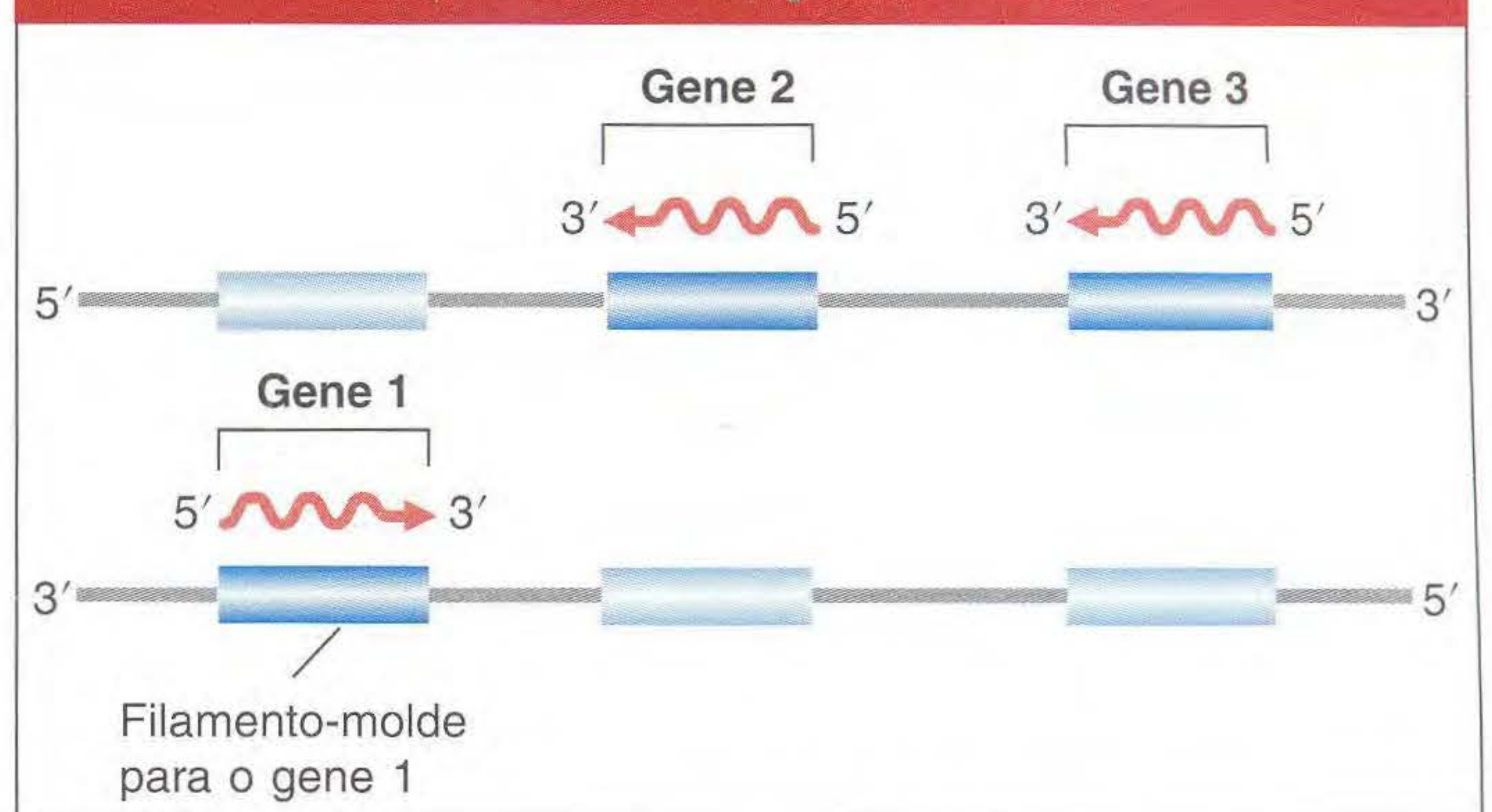


FIG. 8.3 Apenas um filamento de DNA é o molde para a transcrição gênica, mas esse filamento varia com o gene. O sentido da transcrição é sempre o mesmo para qualquer gene e começa da ponta 3' do molde de DNA e da ponta 5' do RNA transcrito. Assim, os genes transcritos em sentidos diferentes usam filamentos opostos do DNA como moldes.

tídeo é posicionado em oposição à sua base complementar pela enzima **RNA polimerase**. Essa enzima liga-se ao DNA e move-se ao longo dele, unindo os ribonucleotídeos alinhados para fazer uma molécula crescente de RNA, como mostrado na Figura 8.4a. Assim, já vimos os dois princípios de complementariedade de bases e a ligação ácido nucleico-proteína em ação (nesse caso, a ligação da RNA polimerase).

Vimos que o RNA tem uma ponta 5' e uma ponta 3'. Durante a síntese, o crescimento do RNA é sempre no sentido 5' para 3'. Em outras palavras, os nucleotídeos são sempre adicionados na ponta crescente 3', como mostrado na Figura 8.4b. Como os filamentos complementares de ácidos nucleicos são orientados em oposição, o fato de que o RNA é sintetizado de 5' para 3' significa que o filamento completo deve ser orientado de 3' para 5'.

À medida que a molécula de RNA polimerase se move ao longo do gene, ela desenrola a dupla de DNA à frente dela e reenrola o DNA que já foi transcrito. À medida que a molécula de RNA progressivamente aumenta, a ponta 5' do RNA é deslocada do molde e a bolha de transcrição fecha-se atrás da polimerase. Várias RNA polimerases, cada uma sintetizando uma molécula de RNA, movem-se ao longo do gene. Os múltiplos filamentos de RNA podem ser vistos vindo como uma única molécula de DNA ao microscópio eletrônico, dando um modo de visualizar o aumento progressivo dos filamentos de RNA (Figura 8.5).

Também já vimos que as bases no transcrito e molde são complementares. Conseqüentemente, a sequência de nucleotídeos no RNA deve ser a mesma que no filamento não-molde do DNA, à exceção de que os T são substituídos por U, como mostrado na Figura 8.6. Quando as sequências de bases do DNA são citadas na literatura científica, a sequência do filamento não-molde é convencionalmente citada, pois essa sequência é a mesma encontrada no RNA. Por esse motivo, o filamento não-molde do DNA é chamado de **filamento codificante**. É extremamente importante ter em mente essa diferenciação quando se discute a transcrição.

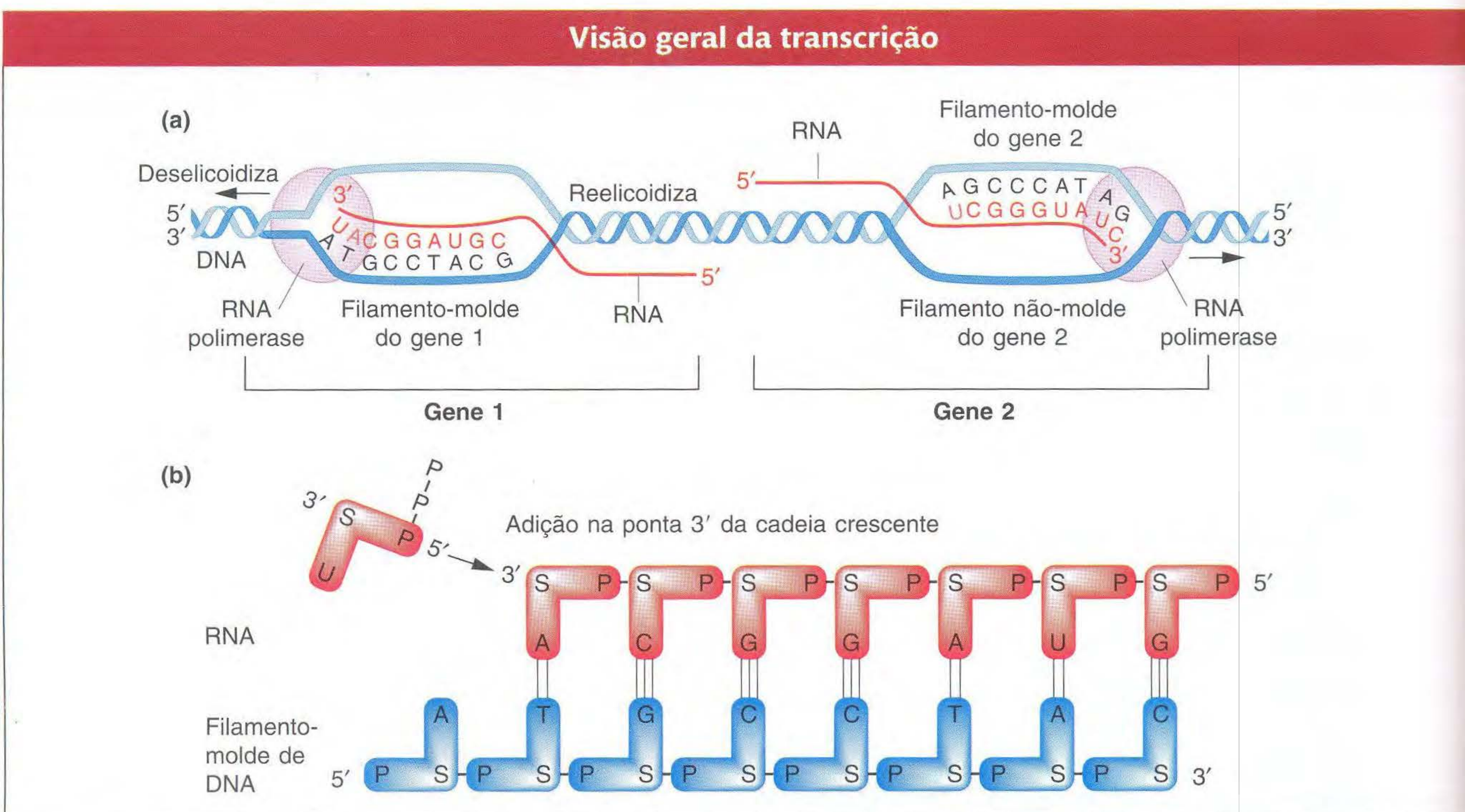


FIG. 8.4 (a) Transcrição dos dois genes em sentidos opostos. Os genes 1 e 2 da Figura 8.3 são mostrados. O gene 1 é transcrito do filamento de baixo. A RNA polimerase migra para a esquerda, lendo o filamento-molde no sentido 3' para 5' e sintetizando RNA em um sentido 5' para 3'. O gene 2 é transcrito no sentido oposto, para a direita, pois o filamento de cima é o molde. À medida que a transcrição continua, a ponta 5' do RNA é deslocada do molde, enquanto a bolha de transcrição se fecha atrás da polimerase. (b) À medida que o gene 1 é transcrito, o grupo fosfato da ponta 5' do nucleotídeo que entra (U) liga-se à ponta 3' da cadeia crescente de RNA. S = açúcar.

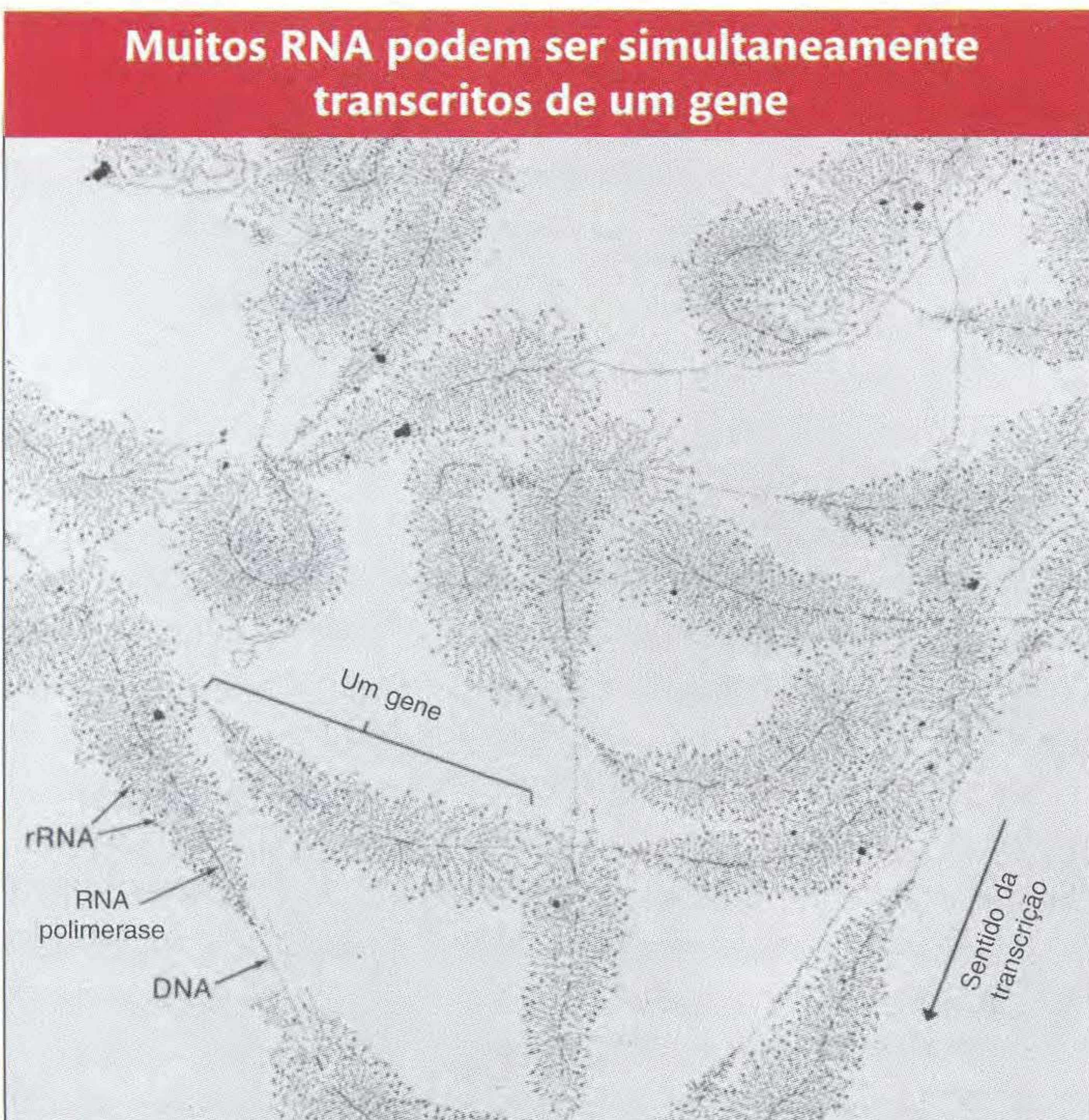


FIG. 8.5 Essa eletromicrografia mostra a transcrição dos genes de RNA ribossômico repetidos em tandem no núcleo do anfíbio *Triturus viridiscens*. Juntamente com cada gene, muitas RNA polimerases são transcritas em uma direção. O crescimento dos RNA transcritos aparece como filamentos que se projetam para fora do arcabouço de DNA. Os transcritos mais curtos estão próximos do início da transcrição; os mais longos estão perto do final do gene. O aspecto de "árvore de Natal" é o resultado. [Foto de O.L. Miller, Jr., e Barbara A. Hamkalo.]

A transcrição é assimétrica: apenas um filamento do DNA de um gene é usado como molde para a transcrição. Esse filamento está na orientação 3' para 5', e o RNA é sintetizado no sentido 5' para 3'.

Estágios da transcrição

A sequência codificante de proteína em um gene é um segmento relativamente pequeno do DNA inserido em uma molécula de DNA muito longa (o cromossomo). Como o segmento apropriado é transcrito em uma molécula unifilar de RNA de tamanho e sequência de nucleotídeos corretos? Como o DNA de um cromossomo é uma unidade contínua, a maquinaria transcrricional deve ser dirigida para o começo de um gene para começar a transcrever no local certo, continuar transcrevendo ao longo do gene e, finalmente, parar de transcrever na outra ponta. Esses três estágios distintos da transcrição são chamados de **iniciação**, **alongamento** e **término**. Embora o processo geral de transcrição seja marcadamente similar em procariontes e eucariontes, existem diferenças importantes. Por esse motivo, seguiremos os três estágios primeiro em procariontes (usando a bactéria intestinal *E. coli* como exemplo) e, então, em eucariontes.

Iniciação em procariontes Como a RNA polimerase encontra o ponto correto de início para a transcrição? Em procariontes, a RNA polimerase geralmente se liga a uma sequência específica de DNA chamada **promotor**, situada

Sequências de DNA e RNA transcrito

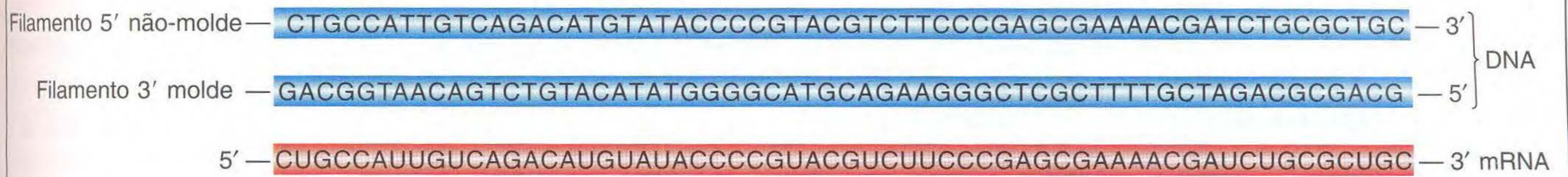


FIG. 8.6 A sequência de mRNA é complementar ao filamento-molde de DNA do qual ele é transcrito e, portanto, corresponde à sequência não-molde (exceto em que o RNA tem U onde o DNA tem T). Essa sequência é do gene para a enzima β -galactosidase.

perto do início da região transcrita. Um promotor é uma parte importante da região reguladora de um gene. Lembre que, como a síntese de um transcrito de RNA começa em sua ponta 5' e continua no sentido 5' para 3', a convenção é desenhar e se referir à orientação do gene no sentido 5' para 3', também. Geralmente, a ponta 5' é desenhada à esquerda e a 3' à direita. Com essa visão, como o promotor deve estar perto da ponta do gene em que começa a transcrição, diz-se que está na ponta 5' do gene; assim, a região promotora também é chamada de região reguladora 5' (Figura 8.7a).

A primeira base transcrita está sempre no mesmo local, chamado de *sítio iniciador*. O promotor é chamado de **antecedente** ao sítio de iniciação porque está situado à frente do sítio de iniciação, na direção oposta à direção da transcrição. Um sítio **posterior** estaria situado depois da região de transcrição. Por convenção, a primeira base do DNA a ser transcrita é numerada +1. As posições dos nucleotídeos ante-

cedentes ao sítio de início são indicadas por sinal negativo (-) e as posteriores por um sinal positivo (+).

A Figura 8.7b mostra as sequências promotoras de sete genes diferentes no genoma de *E. coli*. Como a mesma RNA polimerase se liga a sequências promotoras desses diferentes genes, as similaridades entre os promotores não são surpreendentes. Em particular, duas regiões de grande similaridade aparecem em praticamente cada caso. Essas *regiões* foram chamadas de -35 (menos 35) e -10 porque estão situadas a 35 e 10 pares de bases, respectivamente, antecedentes à primeira base transcrita. Elas são mostradas em amarelo na Figura 8.7b. Como você pode ver, as regiões -35 e -10 de genes diferentes não têm que ser idênticas para desempenhar função similar. Entretanto, é possível chegar a uma sequência de nucleotídeos, chamada **sequência de consenso**, que está de acordo com a maioria das sequências. A sequência de consenso do promotor de *E. coli* é mostrada na parte inferior da Figura 8.7b. Uma holoenzima de RNA polimerase (veja

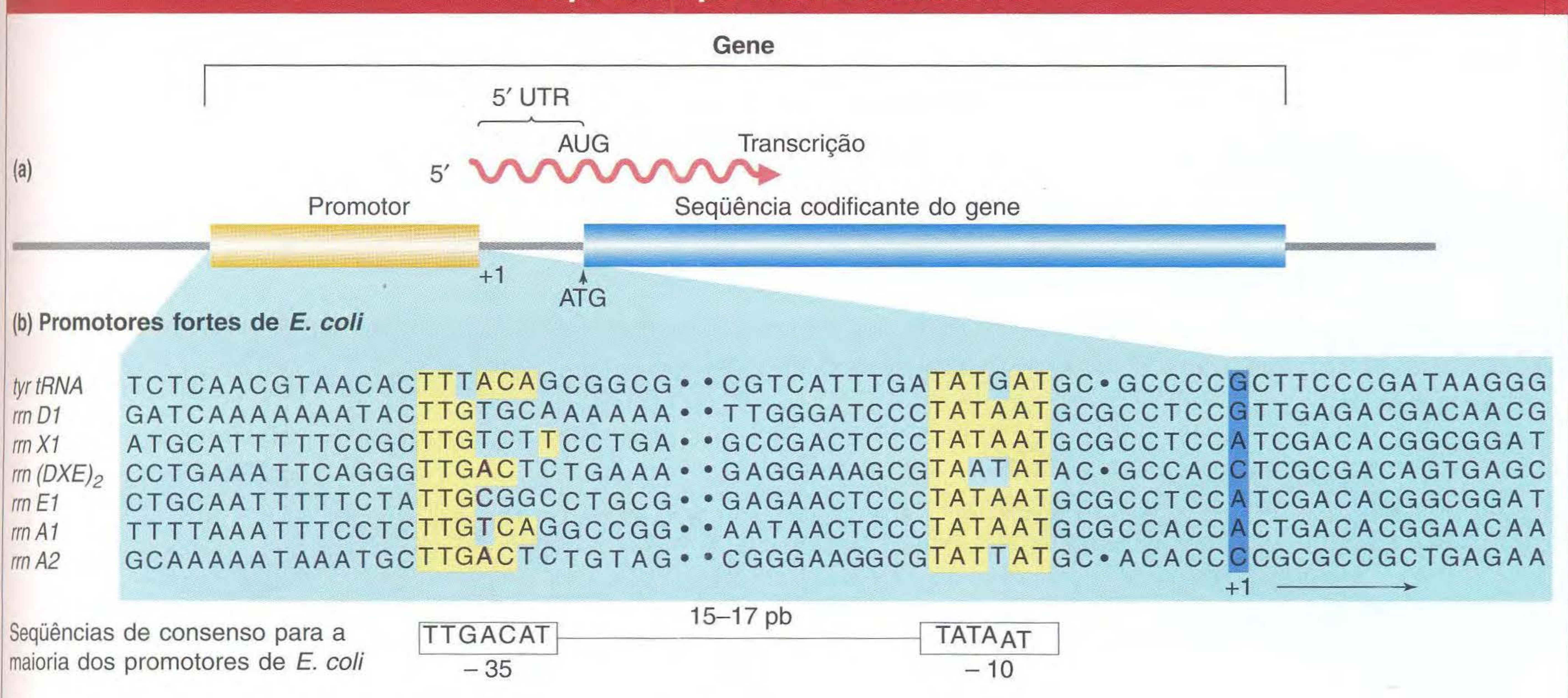
Sequências promotoras em *E. coli*

FIG. 8.7 (a) O promotor fica "antecedente" (para a ponta 5') do ponto de início das sequências codificantes. (b) Os promotores têm regiões de sequências similares como indicado pelo sombreado em amarelo em sete sequências promotoras diferentes em *E. coli*. Os espaços (pontos) são inseridos nas sequências para otimizar o alinhamento das sequências comuns. Os números referem-se ao número de bases antes (-) ou depois (+) do ponto de início da síntese de RNA. A sequência de consenso para a maioria dos promotores de *E. coli* está na parte de baixo. [De H. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, and J. Darnell, *Molecular Cell Biology*, 3rd ed. Copyright 1995 by Scientific American Books, Inc. Todos os direitos reservados. Veja W.R. McClure, *Annu. Rev. Biochem.* 54, 1985, 171, Consensus Sequences.]

o parágrafo seguinte) liga-se ao DNA nesse ponto e, então, desenrola a dupla hélice do DNA, começando a síntese de uma molécula de RNA. Note, na Figura 8.7a, que a parte codificante de proteínas do gene geralmente começa com uma sequência ATG, mas o sítio de iniciação, onde começa a transcrição, geralmente está bem antecedente a essa sequência. A parte intercalar é chamada de **região não-traduzida 5'** (**5' UTR**).

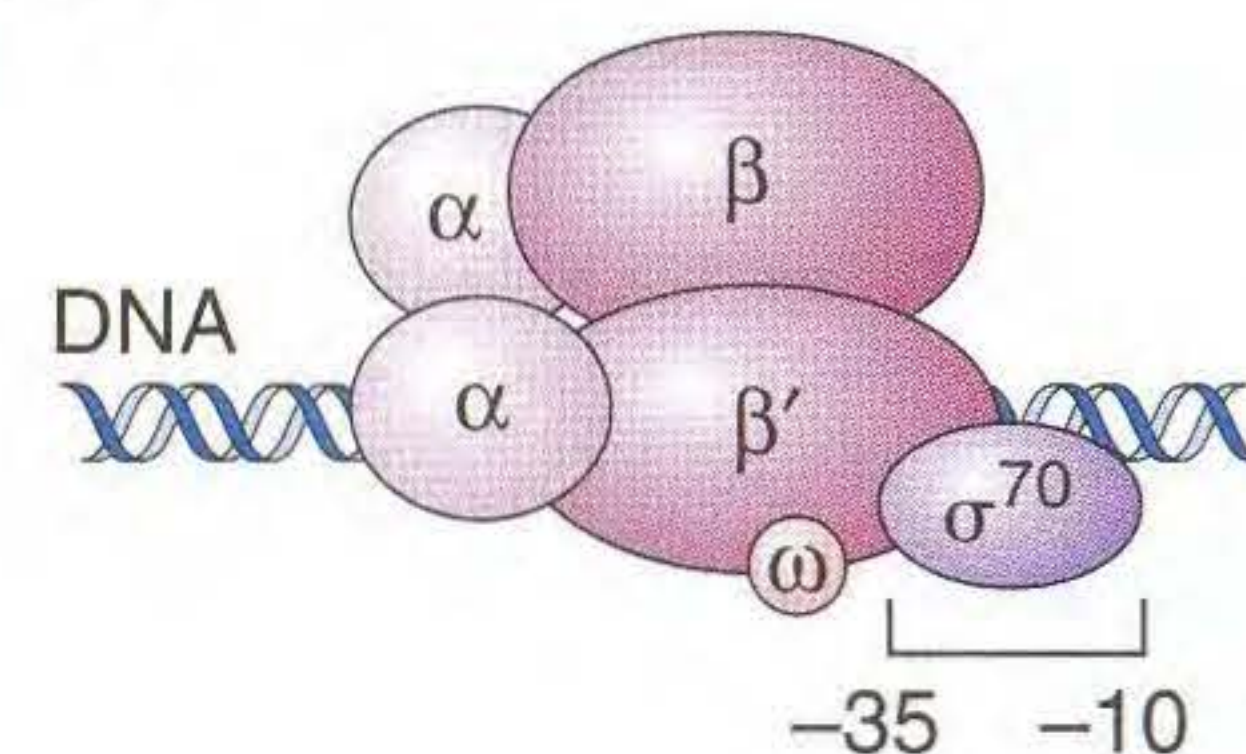
A RNA polimerase bacteriana que procura no DNA uma sequência promotora é chamada de **holoenzima RNA polimerase** (Figura 8.8). Esse complexo multissubunitário é composto de cinco subunidades do cerne da enzima (duas subunidades α , uma β , uma β' e uma ω) mais uma subunidade chamada **fator sigma** (σ). As duas subunidades α ajudam a montar a enzima e promovem interações com proteínas reguladoras, a subunidade β é ativa na catálise, a subunidade β' liga-se ao DNA e a subunidade ω tem papéis na montagem da enzima e na regulação da expressão gênica. A subunidade σ liga-se às regiões -10 e -35, posicionando assim a holoenzima para iniciar corretamente a transcrição no ponto de início (veja Figura 8.8a). A subunidade σ também tem um papel em separar (dissociar) os filamentos de DNA ao redor da região -10, de modo que o cerne da enzima possa se ligar fortemente ao DNA em preparação para a síntese de RNA. Após o cerne da enzima estar ligado, a transcrição começa e a subunidade σ dissocia-se do resto do complexo (veja Figura 8.8b).

E. coli, como a maioria de outras bactérias, tem vários fatores σ diferentes. Um, chamado σ^{70} porque sua massa em quilodaltons é 70, é a subunidade σ primária usada para iniciar a transcrição da grande maioria dos genes de *E. coli*. Outros fatores σ reconhecem sequências promotoras diferentes. Assim, por associação a diferentes fatores σ , a mesma enzima cerne pode reconhecer sequências promotoras diferentes e transcrever conjuntos diferentes de genes.

Alongamento À medida que a RNA polimerase se move ao longo do DNA, desenrola o DNA à frente dela e reenrola o DNA que já foi transcrito. Desse modo, ela mantém uma região unifilamentar de DNA, chamada de **bolha de transcrição**, dentro da qual o filamento-molde é exposto.

Início da transcrição em procariontes

(a) RNA polimerase liga-se a um promotor



(b) Iniciação

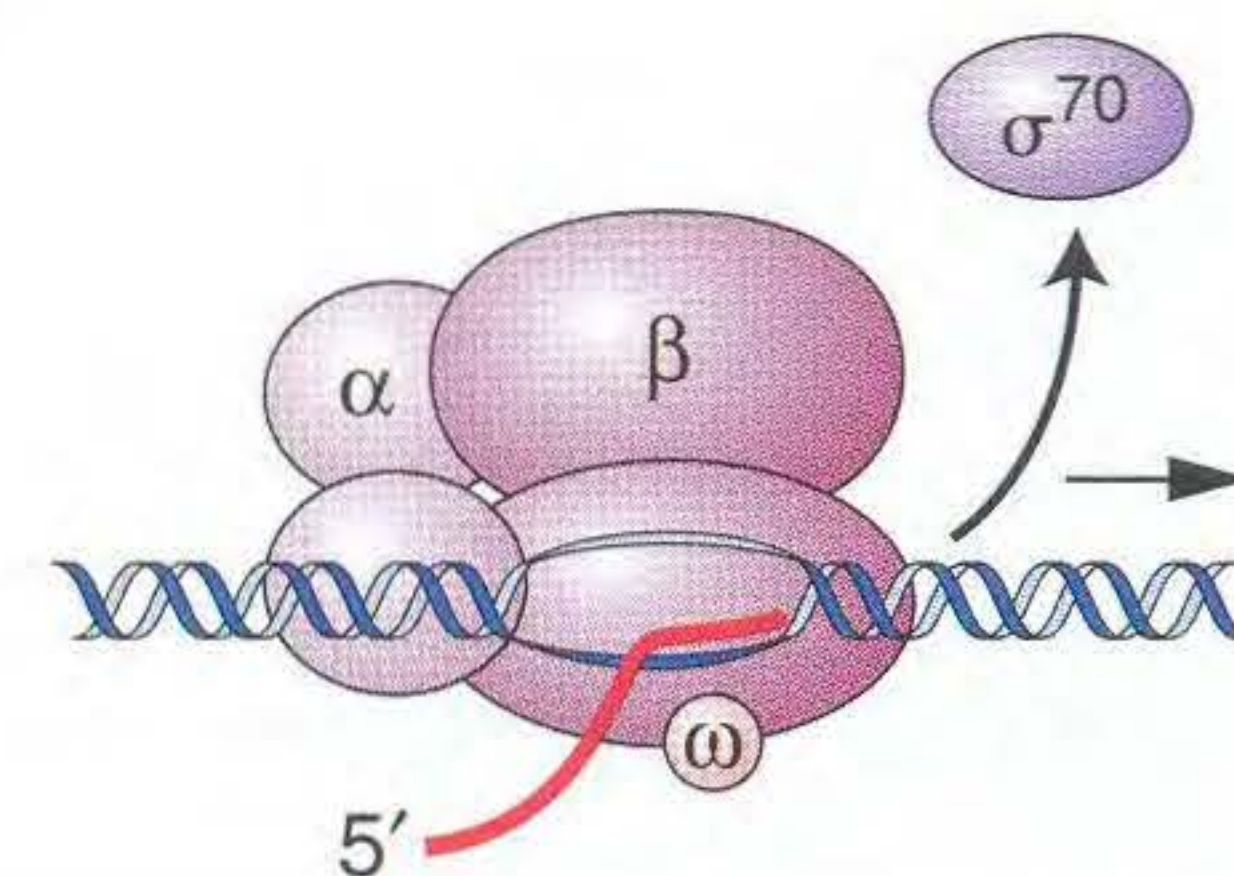
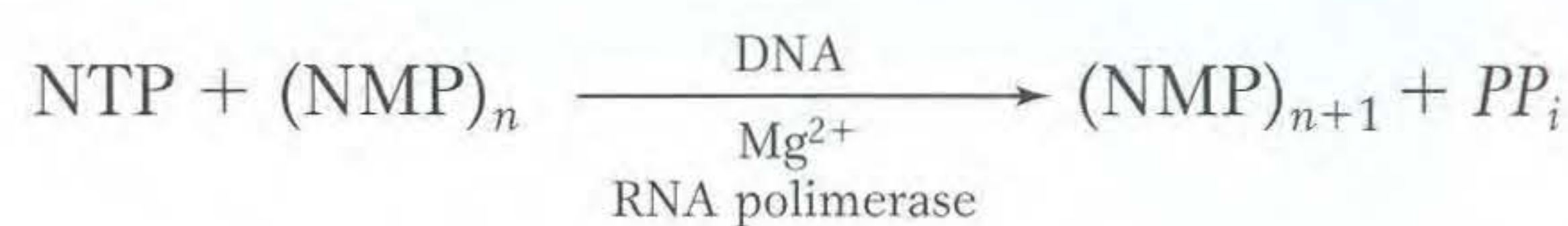


FIG. 8.8 A subunidade σ posiciona a RNA polimerase eucariótica para o início da transcrição. (a) A ligação da subunidade σ às regiões -10 e -35 posiciona a outra subunidade para a iniciação correta. (b) Logo após a síntese do RNA começar, a subunidade σ dissocia-se das outras subunidades, que constinuem a transcrição. [De B.M. Turner, *Chromatin and Gene Regulation*. Copyright 2001 by Blackwell Science Ltd.]

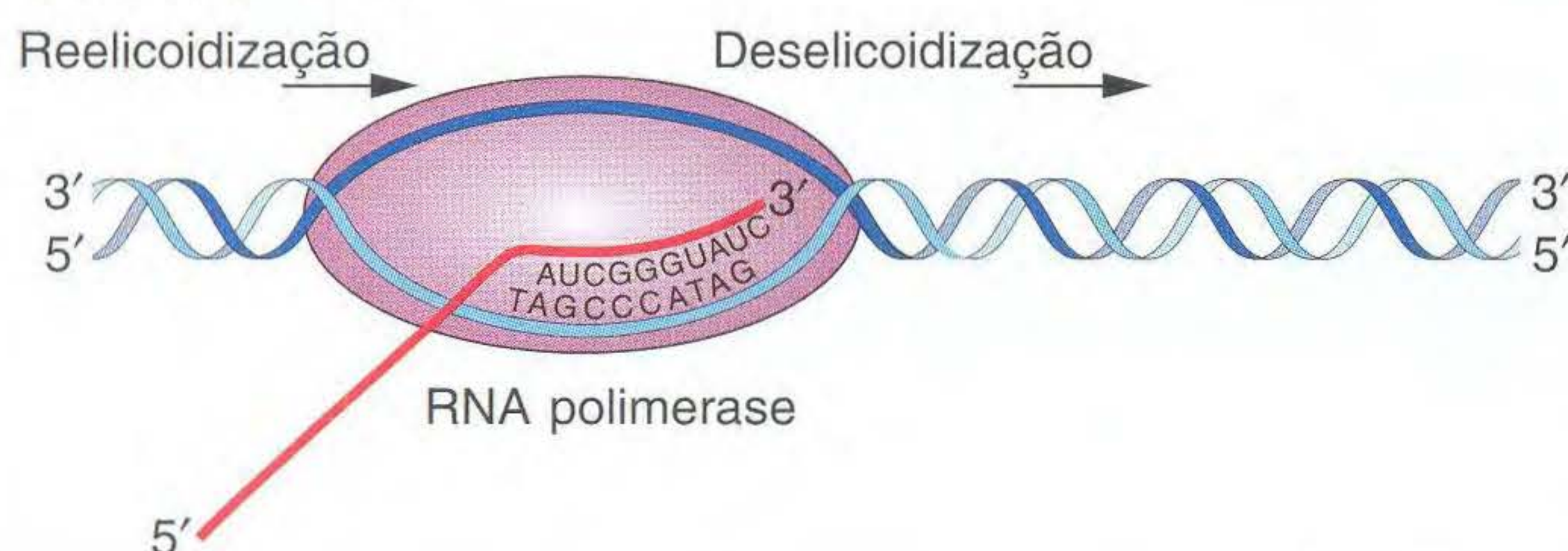
Na bolha, a polimerase monitora a ligação de um ribonucleosídeo trifosfato livre para a base exposta seguinte no molde de DNA, se houver uma complementariedade, e adiciona-o à cadeia. A energia para a adição de um nucleotídeo é derivada da divisão do trifosfato de alta energia e liberação do difosfato inorgânico, de acordo com a seguinte fórmula geral.



A Figura 8.9a mostra um quadro físico do alongamento. Dentro da bolha, os últimos oito ou nove nucleotídeos adicionados à cadeia de RNA formam um híbrido RNA-DNA por pareamento complementar de bases com o filamento-molde.

Alongamento e término da transcrição

(a) Alongamento



(b) Término: mecanismo intrínseco

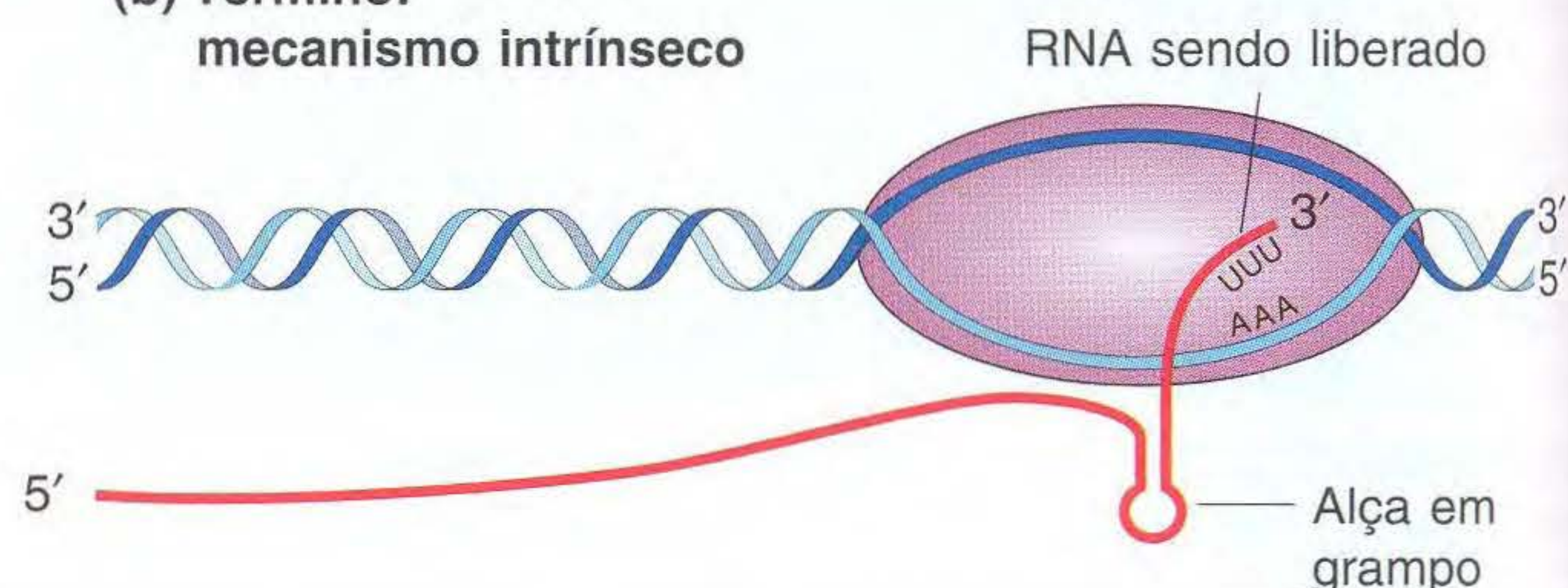


FIG. 8.9 As cinco subunidades da RNA polimerase são mostradas como uma única forma de elipse circundando a bolha de transcrição. (a) **Alongamento**: Síntese de um filamento complementar de RNA à região unifilamentar do filamento-molde de DNA é no sentido 5' para 3'. O DNA desenrolado à frente da RNA polimerase é reenrolado após ela ter sido transcrito. (b) **Término**: O mecanismo intrínseco mostrado aqui é um dos dois modos usados para terminar a síntese de RNA e liberar o RNA transcrito completo e a RNA polimerase do DNA. Nesse caso, a formação de uma alça marca sua liberação. Tanto para o mecanismo intrínseco quanto mediado por ρ , o término primeiro requer a síntese de algumas sequências de RNA.

À medida que a cadeia de RNA aumenta em sua ponta 3', a ponta 5' unifilamentar é liberada da polimerase.

Término A transcrição de um gene individual continua além do segmento codificante de proteína do gene, criando uma **região 3' não-traduzida (3' UTR)** na ponta do transcrito. O alongamento continua até que a RNA polimerase reconheça seqüências especiais de nucleotídeos que atuam como um sinal para o término da cadeia. O encontro com os nucleotídeos de sinal inicia a liberação do RNA nascente e a enzima do molde (Figura 8.9b). Os dois mecanismos principais de término em *E. coli* (e outras bactérias) são chamados de intrínsecos e dependentes de rô.

No mecanismo intrínseco, o término é direto. A seqüência de término contém cerca de 40 pares de bases, terminando em um trecho rico em GC que é seguido por uma fileira de seis ou mais A. Como G e C no molde darão C e G, respectivamente, no transcrito, o RNA nessa região é rico em GC. Essas bases C e G são capazes de formar pontes de hidrogênio complementares uma à outra, resultando em uma alça em grampo (Figura 8.10). Lembre que o par de bases G-C é mais estável que o par A-T, pois tem ponte de hidrogênio em três sítios, enquanto o par A-T (ou A-U) é mantido junto apenas por duas pontes de hidrogênio. As alças em grampo que são amplamente pares G-C são mais estáveis que as alças que são amplamente pares A-U. A alça é seguida de uma fileira de cerca de oito U que correspondem às unidades A do molde de DNA.

Normalmente, no curso do alongamento da transcrição, a RNA polimerase fará uma pausa se o curto híbrido DNA-RNA na bolha de transcrição for fraco e voltará para estabilizar o híbrido. Como nos grampos, a força do híbrido é determinada pelo número relativo de pares de bases G-C comparados com os pares de bases A-U (ou os pares de bases A-T nos híbridos RNA-DNA). No mecanismo intrínseco, a polimerase é tida como fazendo uma pausa após a síntese das U (o que forma um híbrido DNA-RNA fraco). Entretanto, a polimerase retrocedente encontra a alça do grampo. Esse bloqueio gera a liberação do RNA da polimerase e a liberação da polimerase do molde de DNA.

Um sítio de término de transcrição bacteriano

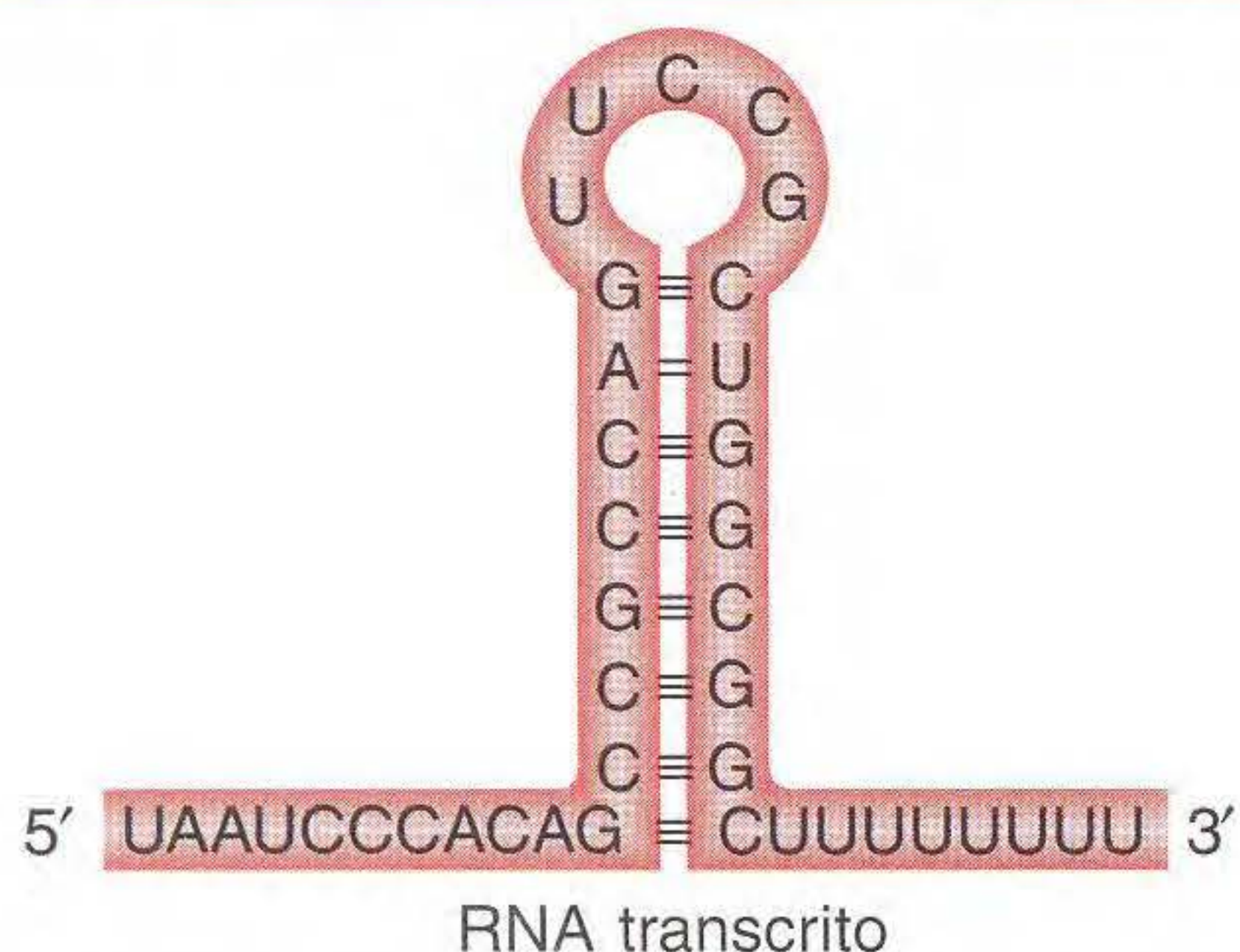


FIG. 8.10 A estrutura de um sítio de término para a RNA polimerase em bactérias. A estrutura em grampo forma-se pelo pareamento complementar de bases com o filamento de RNA rico em GC. A maioria dos pareamentos de bases do RNA é entre G e C, mas há um par A-U.

O segundo tipo de mecanismo de término requer a ajuda de uma proteína chamada fator rô. Essa proteína reconhece os sinais de término para a RNA polimerase. Os RNA com sinais de término dependentes de rô não têm a fileira de unidades U em sua ponta 3', e geralmente não têm alças em grampo. Em vez disso, têm uma seqüência de cerca de 40 a 60 nucleotídeos que é rica em unidades C e pobre em G, e inclui um segmento posterior chamado de sítio *rut* (uso de rô). Rô é um hexâmero consistindo em seis subunidades idênticas que se ligam à cadeia de RNA nascente no sítio *rut*. Esses sítios estão situados logo posteriores às seqüências (lembre que posterior significa 5') em que a RNA polimerase tende a parar. Após a ligação, rô facilita a liberação do RNA da RNA polimerase. Assim, o término dependente de rô envolve a ligação de rô a *rut*, a pausa da polimerase e a dissociação mediada por rô do RNA da RNA polimerase.

8.3 Transcrição em Eucariontes

Como descrito no Capítulo 7, a replicação do DNA em eucariontes, embora mais complicada, é muito similar à replicação do DNA em procariontes. De algum modo, o mesmo pode ser dito da transcrição, pois os eucariontes conservam muitos dos eventos associados à iniciação, alongamento e término em procariontes. A replicação do DNA é mais complexa em eucariontes, em grande parte porque há muito mais DNA para ser copiado. A transcrição é mais complicada em eucariontes por três motivos primários.

1. Os genomas eucarióticos maiores têm muito mais genes a serem reconhecidos e transcritos. Enquanto as bactérias geralmente têm alguns milhares de genes, os eucariontes têm dezenas de milhares de genes. Além disso, há muito mais DNA não-codificante nos eucariontes. O DNA não-codificante origina-se por uma variedade de mecanismos que serão discutidos no Capítulo 14. Assim, muito embora os eucariontes tenham mais genes que os procariontes, seus genes são, em média, mais afastados. Por exemplo, enquanto a densidade de genes (número médio de genes por tamanho de DNA) em *E. coli* é 1 gene por 1.400 pb, esse número cai para 1 gene por 9.000 pb para a mosca-das-frutas *Drosophila*, e é apenas de 1 gene por 100.000 pb para humanos. Essa baixa densidade de genes torna a transcrição, especificamente a etapa inicial, um processo muito mais complicado. Nos genomas de eucariontes multicelulares, encontrar o começo de um gene pode ser como procurar uma agulha no palheiro.

Como você verá, os eucariontes lidam com essa situação de vários modos. Primeiro, dividiram o trabalho da transcrição entre três polimerases diferentes.

- a. A RNA polimerase I transcreve genes de rRNA (excluindo o 5S rRNA).
- b. A RNA polimerase II transcreve todos os genes codificantes de proteína, para os quais o transcrito final é o mRNA, e transcreve alguns snRNA.
- c. A RNA polimerase III transcreve os pequenos RNA funcionais (tais como os genes para tRNA, alguns snRNA, e 5S rRNA).

Nesta seção, enfocaremos nossa atenção na RNA polimerase II.

Segundo, os eucariontes requerem a montagem de muitas proteínas em um promotor antes que a RNA polimerase II possa começar a sintetizar RNA. Algumas dessas proteínas, chamadas de **fatores gerais de transcrição (GTF)**, ligam-se antes que a RNA polimerase II se ligue, enquanto outros se ligam depois. O papel dos GTF e sua interação com a RNA polimerase II serão descritos na seção seguinte, no início da transcrição em eucariontes.

2. Uma diferença significativa entre eucariontes e procariontes é a presença de um núcleo nos eucariontes. Nos procariontes, a informação no RNA é quase imediatamente traduzida em uma cadeia de aminoácidos (polipeptídeos), como veremos no Capítulo 9. Nos eucariontes, a transcrição e a tradução são especialmente separadas, a transcrição ocorre no núcleo e a tradução no citoplasma (Figura 8.11). Nos eucariontes, o RNA é sintetizado no núcleo onde está situado o DNA, sendo exportado do núcleo para o citoplasma para tradução.

Antes que o DNA deixe o núcleo, ele deve ser modificado de vários modos. Essas modificações são coletivamente chamadas de **processamento do RNA**. Para distinguir o RNA antes e depois do processamento, o RNA recém-sintetizado é chamado de **transcrito primário** ou **pré-mRNA**, e o termo **mRNA** é reservado para o transcrito totalmente processado que pode ser exportado para fora do núcleo. Como você verá, a metade 5' do RNA sofre processamento, enquanto a metade de 3' ainda está sendo sintetizada. Assim, a RNA polimerase II deve sintetizar RNA e, simultaneamente, coordenar uma gama diversa de eventos processantes. Por esse motivo, entre outros, a RNA polimerase II é uma enzima multissubunitária mais complicada que a RNA polimerase procariótica. De fato, ela é considerada uma outra máquina molecular. A

coordenação do processamento do RNA e a síntese pela RNA polimerase II serão discutidas na seção de alongamento da transcrição em eucariontes.

3. Finalmente, o molde para a transcrição, o DNA genômico, é organizado em cromatina nos eucariontes (veja Capítulo 2), enquanto está praticamente “nu” nos procariontes. Como você aprenderá no Capítulo 11, algumas estruturas de cromatina podem bloquear o acesso da RNA polimerase ao molde de DNA. Essa característica da cromatina evoluiu em um mecanismo muito sofisticado para regular a expressão gênica. Entretanto, uma discussão sobre a influência da cromatina na habilidade da RNA polimerase II em iniciar a transcrição será posta de lado por enquanto e enfocaremos os eventos que ocorrem *após* a RNA polimerase II ter acesso ao molde de DNA.

Início da transcrição em eucariontes

Como já foi dito antes, a transcrição começa nos procariontes quando a subunidade σ da holoenzima RNA polimerase reconhece as regiões -10 e -35 no promotor de um gene. Após o começo da transcrição, a subunidade σ dissocia-se e o cerne da polimerase continua a sintetizar o RNA dentro da bolha de transcrição que se move ao longo do DNA. Similarmente, em eucariontes, o cerne da RNA polimerase II também não pode reconhecer as seqüências promotoras por si só. Entretanto, diferente das bactérias, nas quais o fator σ é uma parte integral da holoenzima polimerase, os eucariontes requerem GTF para se ligarem às regiões no promotor *antes* da ligação da enzima cerne.

O início da transcrição em eucariontes tem algumas características que são reminiscentes do início da replicação nas origens da replicação. Lembre-se, do Capítulo 7, que as pro-

Transcrição e tradução procariótica e eucariótica comparadas

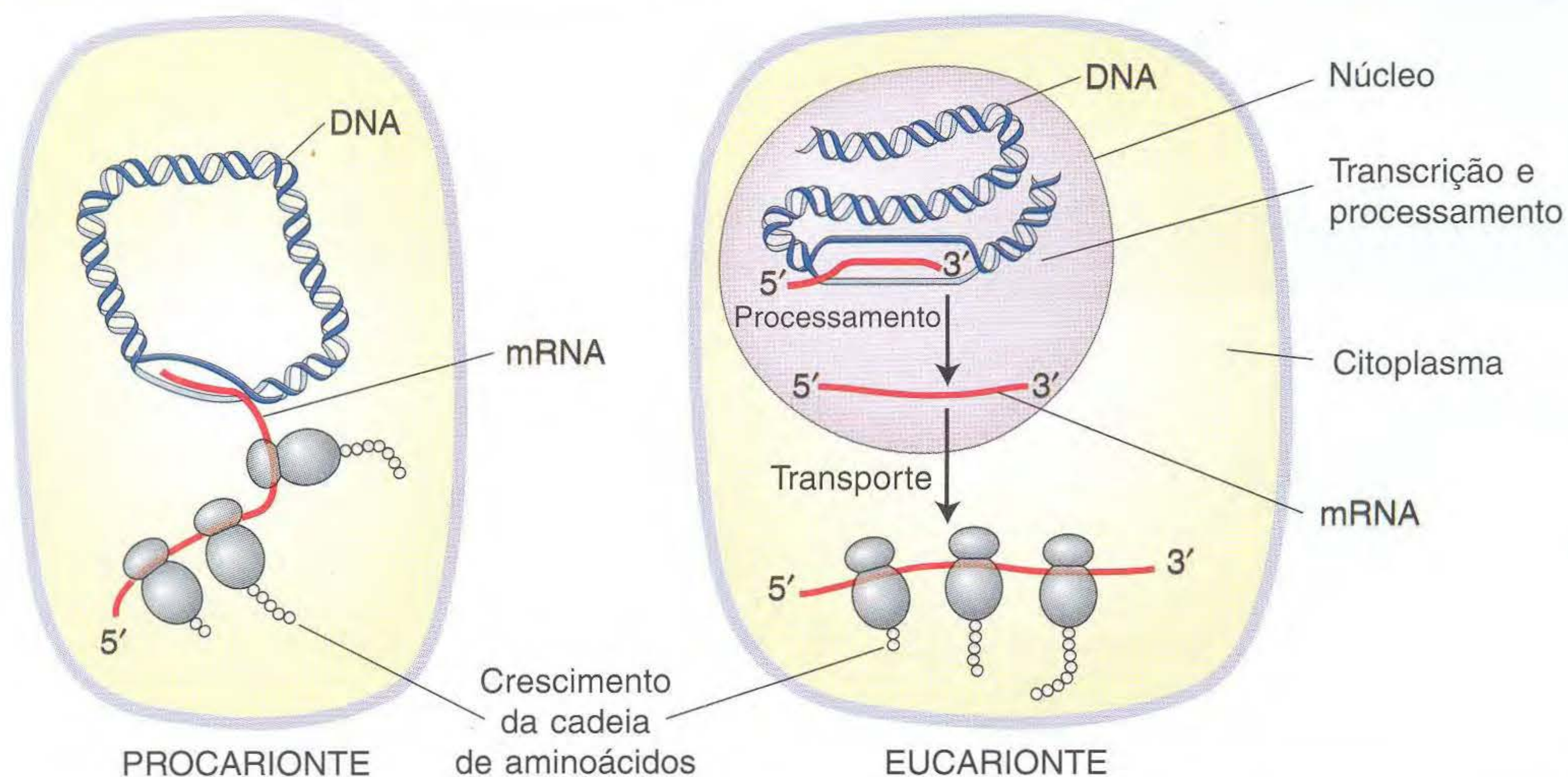


FIG. 8.11 A transcrição e a tradução ocorrem no mesmo compartimento celular em procariontes, mas em compartimentos diferentes nos eucariontes. Além disso, ao contrário dos transcritos de RNA procarióticos, os transcritos eucarióticos sofrem um amplo processamento antes que possam ser traduzidos em proteínas. [De J. Darnell, H. Lodish, and D. Baltimore, *Molecular Cell Biology*, 2nd ed. Copyright 1990 by Scientific American Books, Inc. Todos os direitos reservados.]

teínas que não são parte do replissomo iniciam a montagem da máquina de replicação. O DnaA em *E. coli* e a origem do complexo de reconhecimento (ORC) em leveduras, por exemplo, primeiro reconhecem a origem das seqüências de DNA e ligam-se a estas. Essas proteínas servem como atrativo para as proteínas de replicação, incluindo a DNA polimerase III, por meio de interações proteína-proteína. Similarmente, GTF, que não tomam parte da síntese de RNA, reconhecem as seqüências no promotor e ligam-se a estas ou um a outro, e servem para atrair a RNA polimerase II e posicioná-la no sítio correto para começar a transcrição. Os GTF são designados TFIIA, TFIIB, e assim por diante (para *fator de transcrição da RNA polimerase II*).

Os GTF e o cerne da RNA polimerase II constituem o **complexo de pré-iniciação (PIC)**. Esse complexo é bem grande: ele contém seis GTF, cada um dos quais é um complexo de multiproteína, mais o cerne da RNA polimerase II, que é feito de uma dúzia ou mais de subunidades proteicas. A seqüência de aminoácidos de algumas subunidades do cerne da RNA polimerase II é conservada de leveduras a humanos. Essa conservação pode ser marcadamente demonstrada substituindo-se algumas subunidades da RNA polimerase II de levedura por suas contrapartes humanas para formar uma RNA polimerase *quimérica* (assim denominada devido a uma criatura, na mitologia grega, que soltava fogo e tinha a cabeça de leão, corpo de cabra e cauda de serpente).

Como os promotores procarióticos, os promotores eucarióticos estão situados do lado 5' (posterior) do ponto de início da transcrição. Em um alinhamento das regiões promotoras eucarióticas, a seqüência TATA geralmente pode ser vista situada cerca de 30 pares de bases (-30 pb) do ponto de início da transcrição (Figura 8.12). Essa seqüência, chamada de **TATA boxe**, é o sítio do primeiro evento na transcrição: a ligação da **proteína de ligação a TATA (TBP)**. TBP é parte do complexo TFIID, um dos seis GTF. Quando ligada ao TATA boxe, TBP atrai outros GTF e o cerne da RNA polimerase II para o promotor, formando assim o complexo de pré-iniciação. Após a transcrição ter sido iniciada, a RNA polimerase II dissocia-se da maioria dos GTF para alongar o transcrito primário de RNA. Alguns dos GTF permanecem no promotor para atrair o próximo cerne de RNA polimerase. Desse modo, várias enzimas RNA polimerase II podem sintetizar simultaneamente os transcritos de um único gene.

Como o cerne de RNA polimerase II é capaz de se separar dos GTF e começar a transcrição? Embora os detalhes desse processo ainda estejam sendo desenvolvidos, o que se sabe é que a subunidade β da RNA polimerase II contém uma cauda de proteína, chamada de **domínio da cauda de carboxila (CTD)**, que tem um papel principal. O CTD está estrategicamente situado perto do sítio do qual irá emergir o RNA nascente da polimerase. A fase de iniciação termina e começa a fase de alongamento após CTD ter sido fosforilado por um dos GTF. Essa fosforilação é tida como de algum modo enfraquecendo a ligação da RNA polimerase II a outras proteínas do complexo de pré-iniciação, permitindo o alongamento. O CTD também participa de várias outras fases críticas da síntese e processamento do RNA.

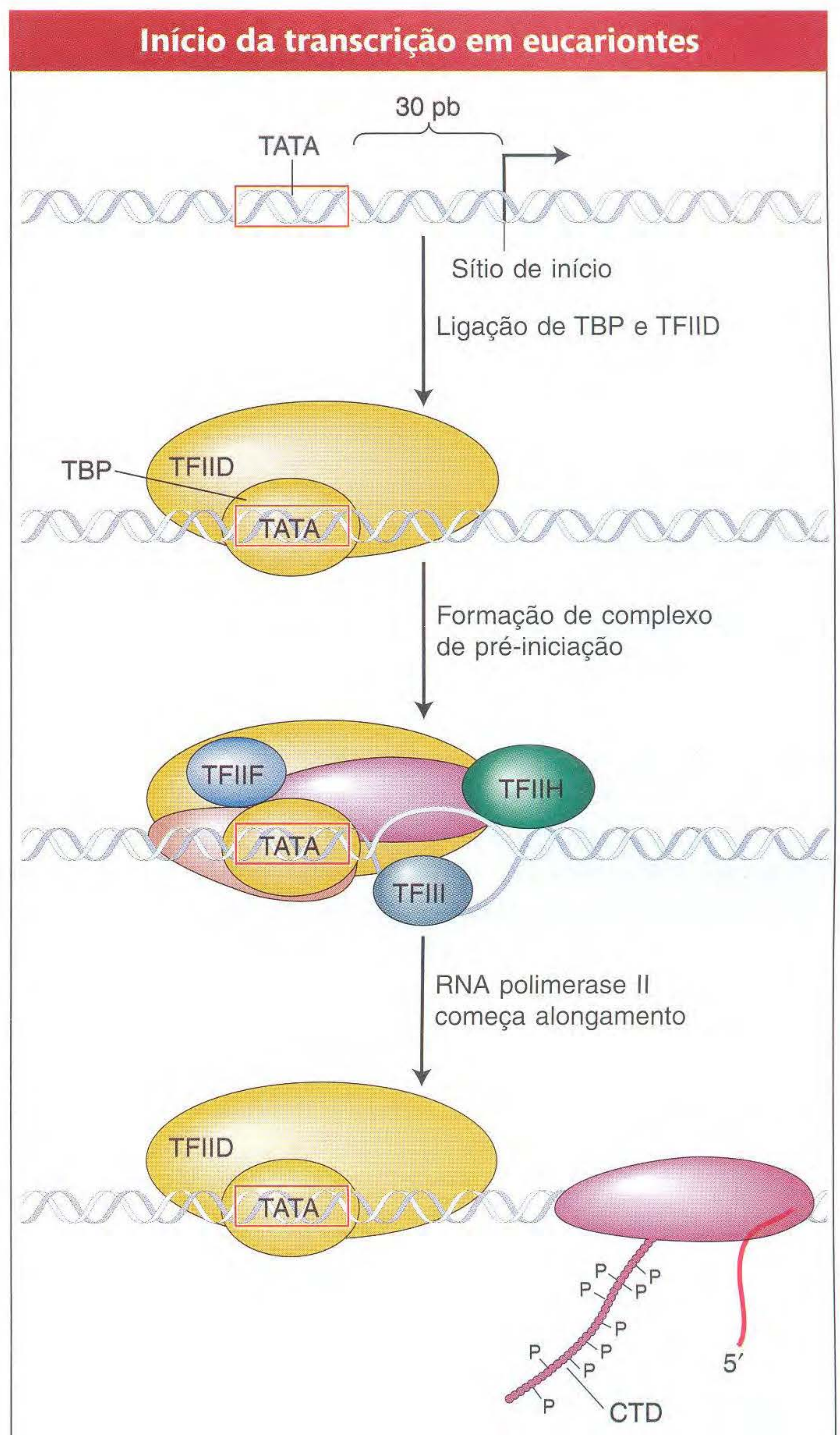


FIG. 8.12 A formação do complexo de pré-iniciação geralmente começa com a ligação da proteína de ligação a TATA (TBP), que então recruta os outros fatores de transcrição (TF) e a RNA polimerase II para o sítio de início da transcrição. A transcrição começa após a fosforilação do domínio da cauda carboxila (CTD) da RNA polimerase II. [De "RNA Polymerase II Holoenzyme and Transcription Factors," *Encyclopedia of Life Sciences*. Copyright 2001 Macmillan Publishing Group Ltd./Nature Publishing Group.]

Os promotores eucarióticos são primeiro reconhecidos pelos fatores gerais de transcrição. A função dos GTF é atrair o cerne da RNA polimerase II de modo que ela é posicionada para começar a síntese de RNA no sítio de início de transcrição.

Alongamento, término e processamento pré-mRNA em eucariontes

O alongamento ocorre dentro da bolha de transcrição essencialmente como descrito para a síntese de RNA procariótico. Entretanto, o RNA nascente tem destinos muito diferentes nos procariontes e eucariontes. Nos procariontes, a tradução

começa na ponta 5' do RNA nascente, enquanto a metade 3' ainda está sendo sintetizada. Em contraste, o RNA dos eucariontes deve sofrer mais processamento antes que possa ser traduzido. Esse processamento inclui (1) a adição de um revestimento (*cap*) na ponta 5', (2) recomposição para eliminar os íntrons e (3) a adição de uma cauda 3' dos nucleotídeos adenina (poliadenilação).

Como a replicação do DNA, a síntese e processamento do pré-mRNA em mRNA requer que muitas etapas sejam feitas com rapidez e precisão. No início, considerava-se que a maioria do processamento de pré-mRNA eucariótico ocorresse após a síntese de RNA estar completa. O processamento após a síntese do RNA estar completa é dito **pós-transcricional**. Entretanto, as evidências experimentais agora indicam que o processamento de fato ocorre durante a síntese de RNA; ele é **co-transcricional**. Portanto, o RNA parcialmente sintetizado (nascente) está sofrendo reações de processamento à medida que emerge do complexo RNA polimerase II.

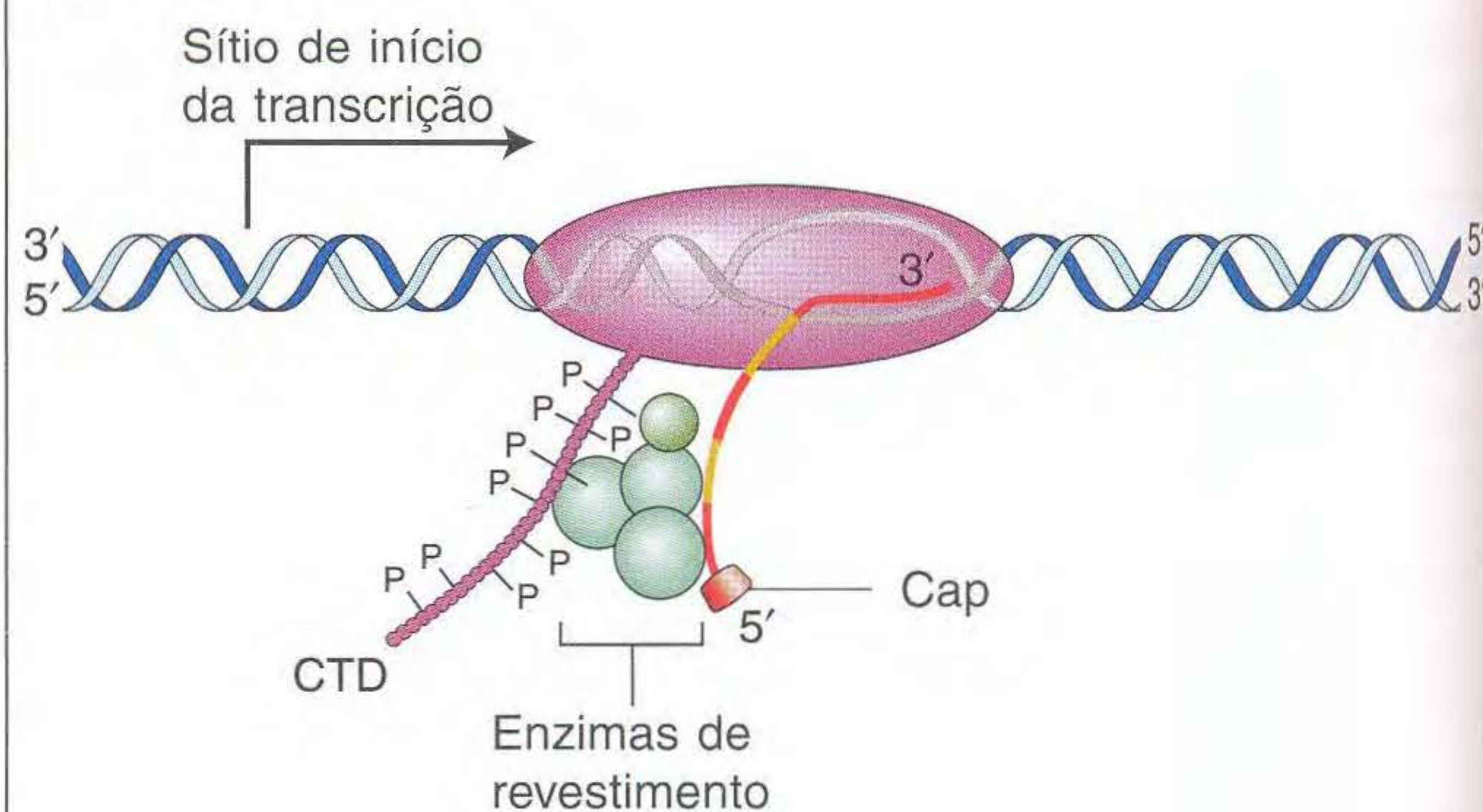
O CTD da RNA polimerase II eucariótica tem um papel central em coordenar todos os eventos de processamento. O CTD é composto de muitas repetições de uma sequência de sete aminoácidos. Essas repetições servem como sítios de ligação para algumas das enzimas e outras proteínas que são necessárias para o revestimento do RNA, recomposição e clivagem seguidos de poliadenilação. O CTD está situado perto do sítio de onde emerge da polimerase o RNA nascente, e, assim, é um local ideal para orquestrar a ligação e liberação das proteínas necessárias para processar o RNA nascente transcrito enquanto a síntese do RNA continua. Nas várias fases do processamento, os aminoácidos de CTD são reversivelmente modificados, geralmente pela adição e remoção de grupos fosfato (chamadas de fosforilação e desfosforilação, respectivamente). O estado de fosforilação de CTD determina que proteínas de processamento podem se ligar. Desse modo, CTD determina a tarefa a ser desempenhada no RNA à medida que ele emerge da polimerase. Os eventos de processamento e o papel de CTD em executá-los são mostrados na Figura 8.13 e considerados no texto.

Processamento das pontas 5' e 3' A Figura 8.13a mostra o processamento da ponta 5' do transcrito de um gene codificante de proteína. Quando o RNA nascente emerge de uma RNA polimerase II, uma estrutura especial, chamada **revestimento** (*cap*), é adicionada à ponta 5' por várias proteínas que interagem com CTD. O revestimento consiste em uma 7-metilguanosina ligada ao transcrito por três grupos fosfato. O revestimento tem duas funções: primeira, ele protege o RNA da degradação, uma etapa importante, considerando que um mRNA eucariótico tem uma longa jornada antes de ser traduzido; segunda, como você verá no Capítulo 9, o revestimento é necessário para a tradução do mRNA.

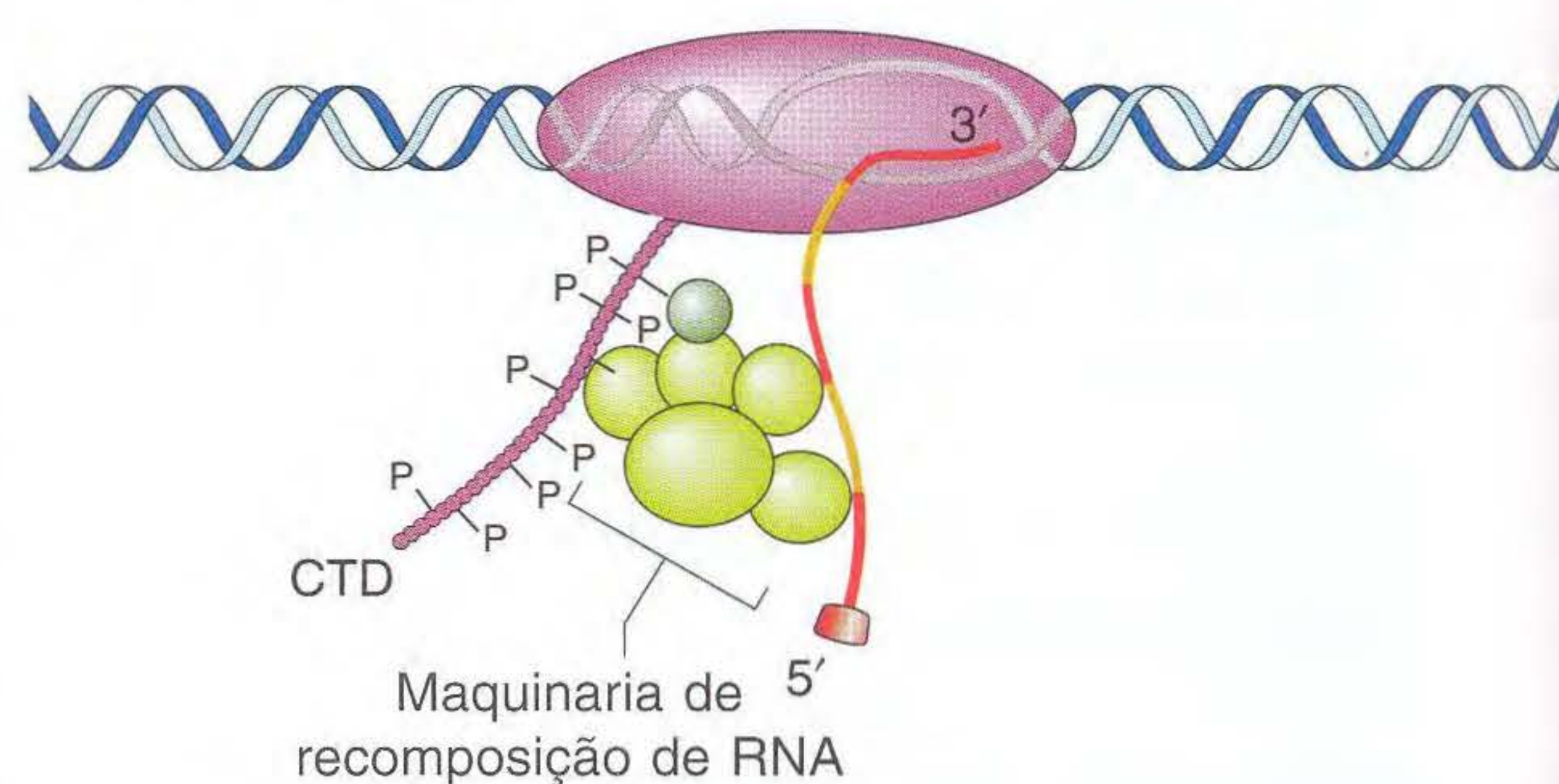
O alongamento do RNA continua até que seja reconhecida a sequência AAUAAA ou AUUAAA perto da ponta 3' por uma enzima que corta a ponta do RNA aproximadamente 20 bases depois. É adicionado a essa ponta um trecho de 150 a 200 nucleotídeos adenina chamados **cauda poli(A)** (veja Figura 8.13c). Assim, a sequência AAUAAA do mRNA dos genes codificantes de proteína é chamada de *senal de poliadenilação*.

Processamento co-transcricional do RNA

(a) Revestimento



(b) Recomposição



(c) Clivagem e poliadenilação

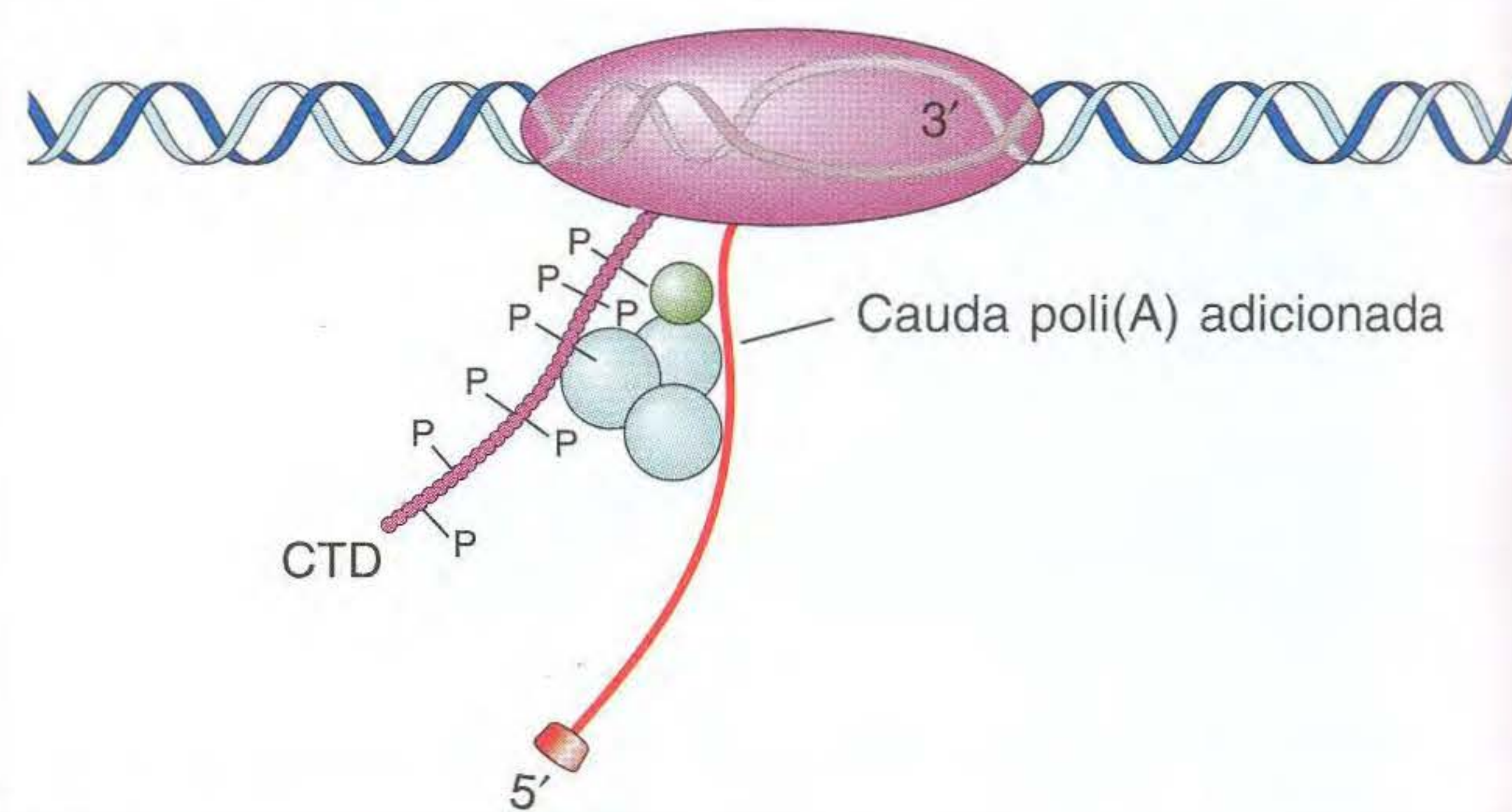


FIG. 8.13 O processamento co-transcricional de RNA é coordenado pelo domínio da cauda carboxila (CTD) da subunidade β da RNA polimerase II. A fosforilação reversível dos aminoácidos de CTD (indicada por P) cria sítios de ligação para as diferentes enzimas de processamento e fatores necessários para (a) revestimento, (b) recomposição e (c) clivagem e poliadenilação. [Partes (a) e (b) de R.I. Drapkin and D.F. Reinberg, "RNA Synthesis," *Encyclopedia of Life Sciences*. Copyright 2002, Macmillan Publishing Group Ltd./Nature Publishing Group.]

Recomposição do RNA, a remoção dos íntrons Em 1977, um estudo científico surgiu com o título "Um incrível arranjo de sequência na ponta 5' do RNA mensageiro de adenovírus 2." Os cientistas geralmente não têm muito reconhecimento, pelo menos em suas publicações, e o uso da palavra "incrível" indicava que algo verdadeiramente inesperado tinha sido encontrado. Os laboratórios de Richard Roberts e Phillip Sharp tinham, independentemente, descoberto que a

informação codificada por genes eucarióticos (em seu caso, o gene de um vírus que infecta células eucarióticas) pode ser fragmentada em pedaços de dois tipos: éxons e íntrons. Como já foi dito, os pedaços que codificam partes das proteínas são os *éxons*, e os pedaços que separam os éxons são os *íntrons*. Os íntrons estão presentes não apenas nos genes codificantes de proteínas, mas também em alguns rRNA e, até mesmo, tRNA. Os íntrons são removidos do transcrito primário enquanto o RNA ainda está sendo transcrito e após o revestimento ter sido adicionado, mas antes de o transcrito ser transportado para o citoplasma. A remoção dos íntrons e a união dos éxons é chamada de **recomposição (*splicing*)**, pois é remanescente do modo pelo qual um videotape ou um filme pode ser cortado e reunido para deletar um segmento específico. A recomposição junta as regiões codificantes, ou éxons, de modo que o mRNA agora contenha a sequência codificante que é completamente colinear à proteína que ela codifica.

O número e tamanho dos íntrons variam de gene para gene e de espécie para espécie. Por exemplo, apenas 200 dos 6.300 genes em leveduras têm íntrons, enquanto os genes típicos nos mamíferos, incluindo os humanos, têm muitos. O tamanho médio de um íntron de mamífero é de cerca de 2.000 nucleotídeos, e um éxon médio tem cerca de 200 nucleotídeos. Assim, a maior percentagem de DNA nos mamíferos codifica íntrons, e não éxons. Um exemplo extremo é o gene humano da distrofia muscular de Duchenne. Esse gene tem 79 éxons e 78 íntrons em seus 2,5 milhões de pares de bases.

Quando recompostos, seus 79 éxons produzem um mRNA de 14.000 nucleotídeos, o que significa que os íntrons correspondem à grande maioria dos 2,5 milhões de pares de bases.

Recomposição alternativa Neste ponto, você pode estar imaginando sobre a utilidade de ter genes organizados em éxons e íntrons. Lembre que este capítulo começou com uma discussão sobre o número de genes no genoma humano. Esse número (± 25.000 genes) é apenas o dobro do número de genes no nematelminto, embora o espectro de proteínas humanas (chamado de **proteoma**, veja Capítulo 9) excede a 100.000. Essas proteínas superam o número de genes de tal modo que indicam que um gene pode codificar a informação para mais de uma proteína. Um modo pelo qual um gene pode codificar várias proteínas é por um processo chamado de **recomposição alternativa**. Por esse processo, diferentes mRNA e, subsequentemente, diferentes proteínas são produzidos pelo mesmo transcrito primário, recompondo diferentes combinações de éxons. Por motivos ainda desconhecidos, a proporção de genes alternativamente recompostos varia de espécie a espécie. Embora a recomposição alternativa seja rara em plantas, mais de 70% dos genes humanos são alternativamente recompostos. Muitas mutações com sérias consequências para o organismo são devidas a defeitos de recomposição.

As consequências da recomposição alternativa na estrutura e função da proteína serão apresentadas mais adiante neste livro. Por agora é suficiente dizer que as proteínas

Padrões complexos de recomposição eucariótica de mRNA

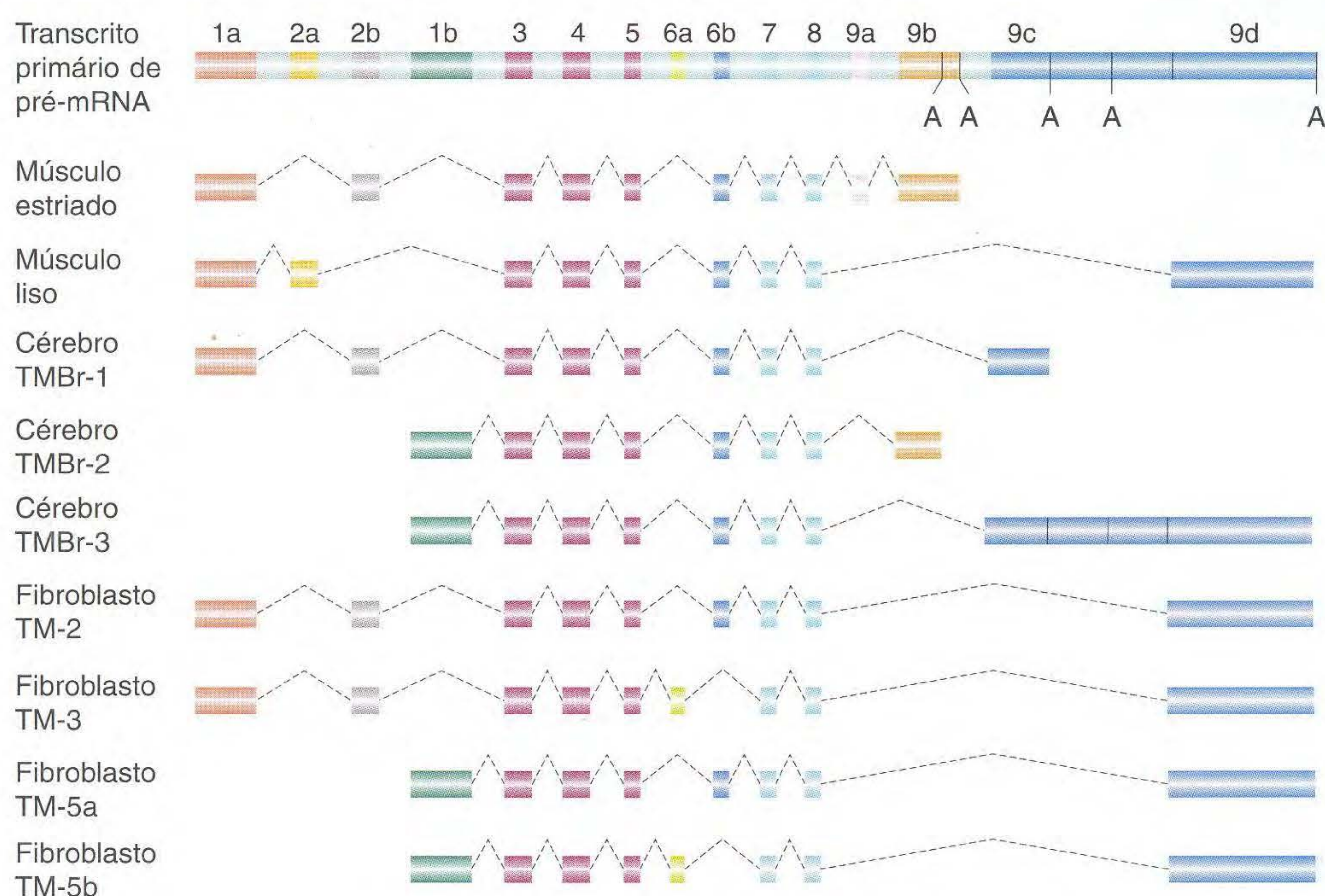


FIG. 8.14 O pré-mRNA transcrito do gene de α -tropomiosina de rato é alternativamente recomposto em diferentes tipos celulares. Os boxes verde-claros representam íntrons; as outras cores representam éxons. Os sinais de poliadenilação são indicados por um A. As linhas tracejadas nos mRNA maduros indicam regiões que foram removidas pela recomposição. TM, tropomiosina. [De J.P. Lees et al., *Mol. Cell Biol.* 10, 1990, 1729-1742.]

produzidas pela recomposição alternativa estão geralmente relacionadas (porque, geralmente, contêm subgrupos dos mesmos éxons do transcrito primário) e elas são geralmente usadas em tipos diferentes de células ou em estágios diferentes do desenvolvimento. A Figura 8.14 mostra a miríade de combinações produzidas pela recomposição alternativa do RNA primário transcrito do gene de α -tropomiosina. O mecanismo de recomposição é considerado na seção seguinte.

O pré-mRNA eucariótico é amplamente processado pelo revestimento 5', poliadenilação de 3' e remoção de íntrons e recomposição de éxons antes que ele possa ser transportado como mRNA para o citoplasma para tradução em proteínas. Um gene pode codificar mais de um polipeptídeo quando seu pré-mRNA é alternativamente recomposto.

► 8.4 RNA Funcionais

Como o RNA é uma molécula tão versátil, ele participa de uma variedade de processos celulares. No Capítulo 9, você aprenderá mais sobre o papel dos **RNA funcionais** como componentes importantes do ribossomo, a máquina biológica que é o sítio da síntese de proteínas. Nesta seção, você verá que os RNA funcionais também têm papéis proeminentes tanto no processamento do mRNA quanto na regulação de seu nível na célula.

Pequenos RNA nucleares (snRNA): o mecanismo de recomposição dos éxons

Após a descoberta dos éxons e íntrons, os cientistas enfocaram sua atenção no mecanismo de recomposição do RNA. Como os íntrons devem ser removidos com precisão e os éxons corretamente unidos, o primeiro enfoque foi comparar as seqüências dos pré-mRNA quanto a indícios de como os íntrons e éxons são reconhecidos. A Figura 8.15 mostra as junções éxon-íntron dos pré-mRNA. Essas junções são os sítios nos quais ocorrem as reações de recomposição. Nessas junções alguns nucleotídeos específicos foram encontrados sendo quase idênticos entre os genes e entre as espécies. Elas têm sido altamente conservadas porque participam das reações de recomposição. Cada íntron é cortado em cada ponta, e essas pontas de íntron geralmente têm GU na ponta 5' e AG

nas ponta 3' (a **regra GU-AG**). Outro sítio invariante é uma A entre 15 e 45 nucleotídeos antecedentes ao sítio de corte 3'. Os nucleotídeos que flanqueiam os altamente conservados também são conservados, mas em menor grau. A existência de seqüências de nucleotídeos nas junções de corte sugeriu que deve haver uma maquinaria que reconheça essas seqüências e faça a recomposição. Como em geral ocorre na pesquisa científica, a maquinaria de recomposição foi encontrada por acaso e o mecanismo de recomposição era totalmente inesperado.

Um achado no laboratório de Joan Steitz levou à descoberta dos componentes da maquinaria de recomposição. Os pacientes com uma variedade de doenças auto-imunes, incluindo o lúpus eritematoso sistêmico, produzem anticorpos contra suas próprias proteínas. No curso da análise de amostras de sangue dos pacientes com lúpus, Joan Steitz *et al.* identificaram anticorpos que podiam ligar-se a um grande complexo molecular de pequenos RNA e proteínas. Como esse complexo riboproteico foi localizado no núcleo, os componentes de RNA foram chamados de *pequenos RNA nucleares*. Os snRNA foram vistos sendo complementares às seqüências de consenso nas junções de recomposição, levando os cientistas a supor um papel para os snRNA na reação de recomposição. Os nucleotídeos conservados no transcrito hoje são reconhecidos por cinco pequenas ribonucleoproteínas nucleares (snRNP), que são complexos de proteína e um dos cinco snRNA (U1, U2, U4, U5 e U6). Esses snRNP e mais de 100 proteínas adicionais são parte do spliceossomo, a grande máquina biológica que remove os íntrons e une os éxons. Os componentes do spliceossomo interagem com CTD, como sugerido na Figura 8.13b.

Esses componentes do spliceossomo ligam-se a seqüências de íntron e éxon, como mostrado na Figura 8.16. Os snRNP ajudam a alinhar os sítios de corte tanto no final de um íntron, formando pontes de hidrogênio com o íntron conservado quanto com as seqüências de éxon. Então os snRNP recrutam outros complexos e formam o spliceossomo, que catalisa a remoção do íntron por duas etapas consecutivas de recomposição (veja Figura 8.16). A primeira etapa liga uma ponta de um íntron à adenina interna conservada, formando uma estrutura que tem a forma de um laço de *cowboy*. A segunda etapa libera o laço e junta os dois éxons adjacentes. A Figura 8.17 representa a química subjacente à remoção do íntron e recomposição dos éxons. Quimicamente, as duas etapas são reações de transesterificação entre os nucle-

Seqüências conservadas relacionadas à remoção de íntrons

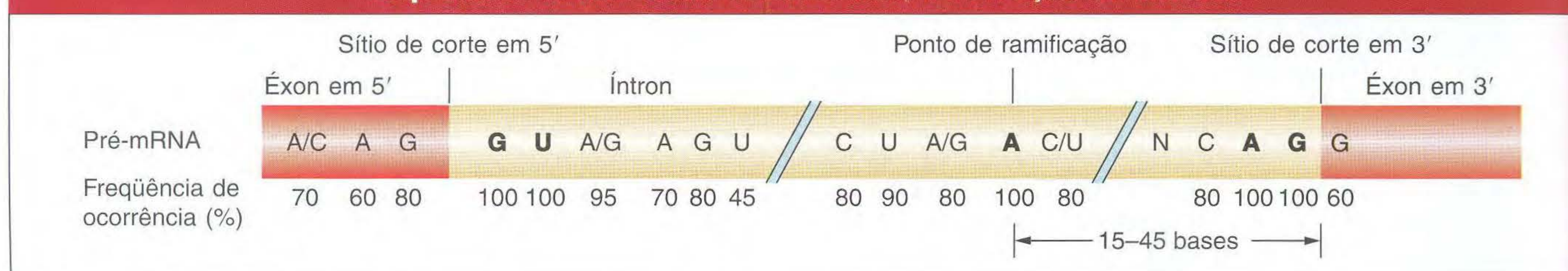


FIG. 8.15 Seqüências de nucleotídeos conservadas estão presentes nas junções de íntrons e éxons. Os números abaixo dos nucleotídeos indicam a percentagem de similaridade entre os organismos. As unidades G e U são de particular importância na ponta 5', as unidades A e G na ponta 3' e a unidade A marcada como "ponto de ramificação" (veja Figura 8.17 para uma visão da estrutura de ramificação). N representa qualquer base.

Montagem e funcionamento do spliceossomo

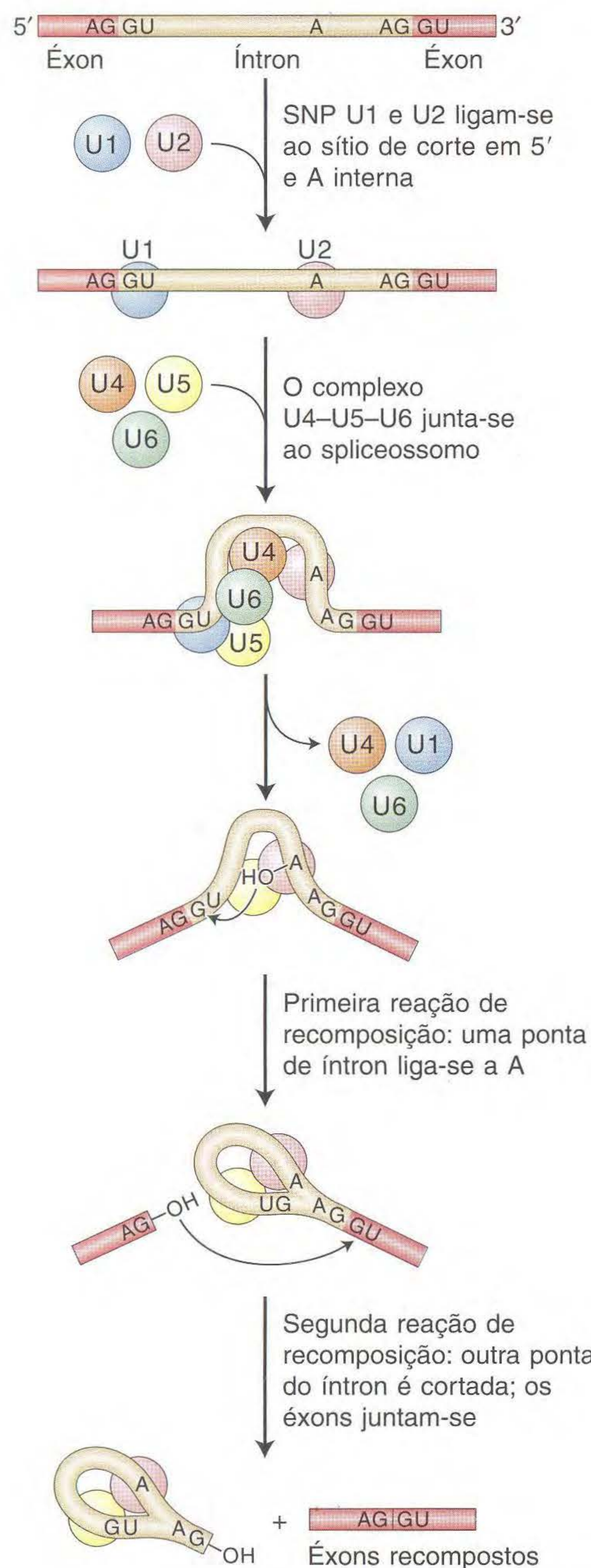


FIG. 8.16 O spliceossomo é composto de vários snRNP que se ligam sequencialmente ao RNA, nas posições mais ou menos como as mostradas. O alinhamento dos snRNP resulta das pontes de hidrogênio de suas moléculas de snRNA com as seqüências complementares do íntron. Desse modo, os reagentes são apropriadamente alinhados e podem ocorrer duas reações de recomposição. A química dessas reações pode ser vista em mais detalhes na Figura 8.17.

tídeos conservados. Os grupos hidroxila nas posições 2' e 3' dos ribonucleotídeos são participantes fundamentais da reação.

Auto-recomposição de íntrons e o mundo de RNA

Dois casos excepcionais de processamento de RNA levaram a uma descoberta considerada por alguns como sendo tão

Reações na recomposição de éxons

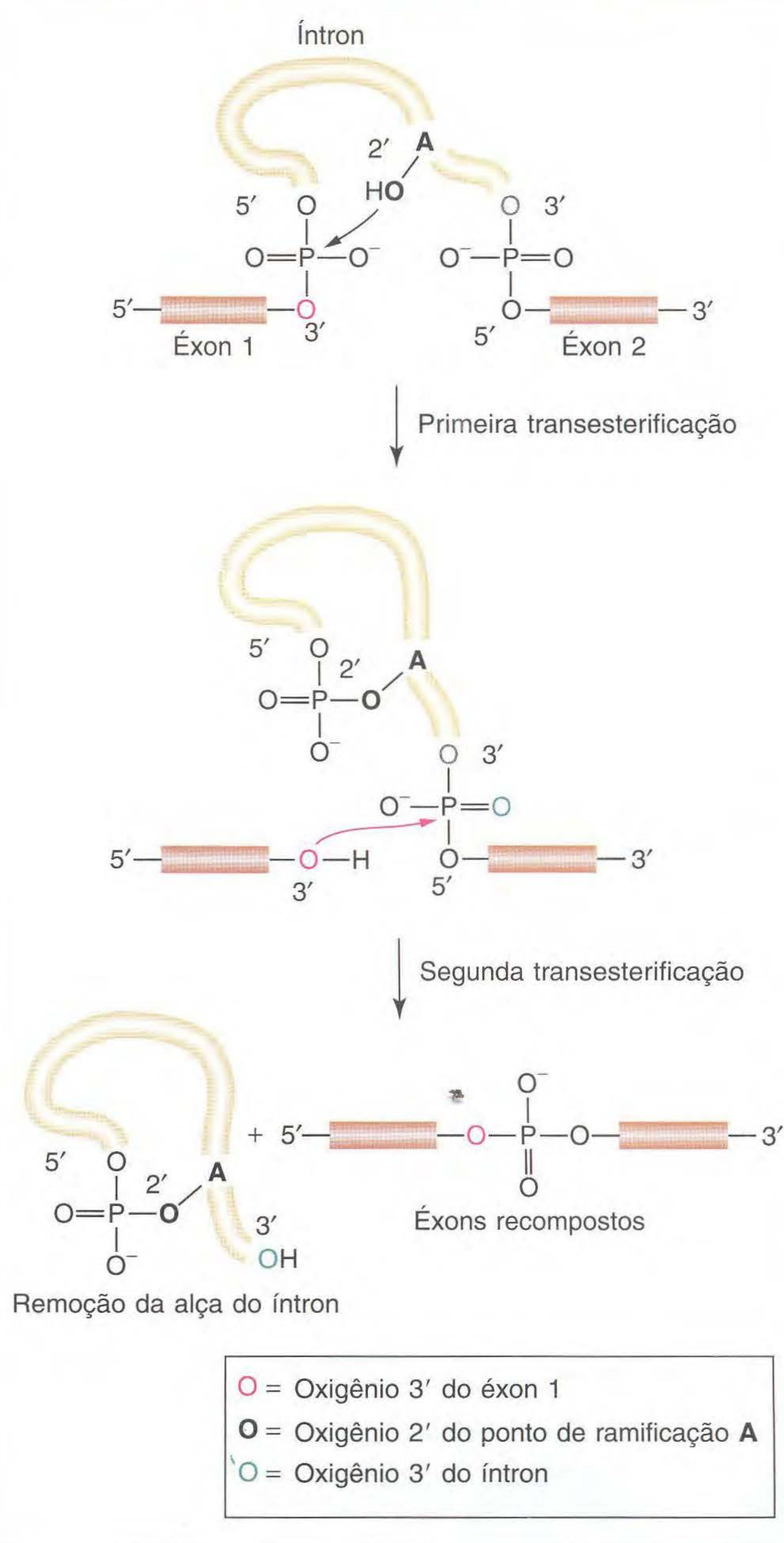


FIG. 8.17 Duas reações de transesterificação ocorrem na recomposição do RNA: primeiro, para juntar a ponta 5' doadora com o ponto de ramificação interno (primeira reação na Figura 8.16) e, segundo, juntar os dois éxons (segunda reação na Figura 8.16). [De H. Lodish, A. Berk, S.O. Zipurski, P. Matsudaira, D. Baltimore, and J. Darnell, *Molecular Cell Biology*, 4th ed. Copyright 2000 by W.H. Freeman and Company.]

importante quanto a estrutura da dupla hélice do DNA. Em 1981, Tom Cech *et al.* relataram que, em um tubo de ensaio, o transcrito primário de um rRNA do protozoário ciliado *Tetrahymena* podia remover um íntron de 413 nucleotídeos de *si mesmo* sem a adição de nenhuma proteína (Figura 8.18). Subseqüentemente, foram demonstrados outros íntrons com essa propriedade, os quais ficaram conhecidos como **íntrons de auto-remoção**. Alguns anos antes, enquanto estudava o processamento do tRNA em bactérias, Sidney Altman identi-

Reação de auto-recomposição

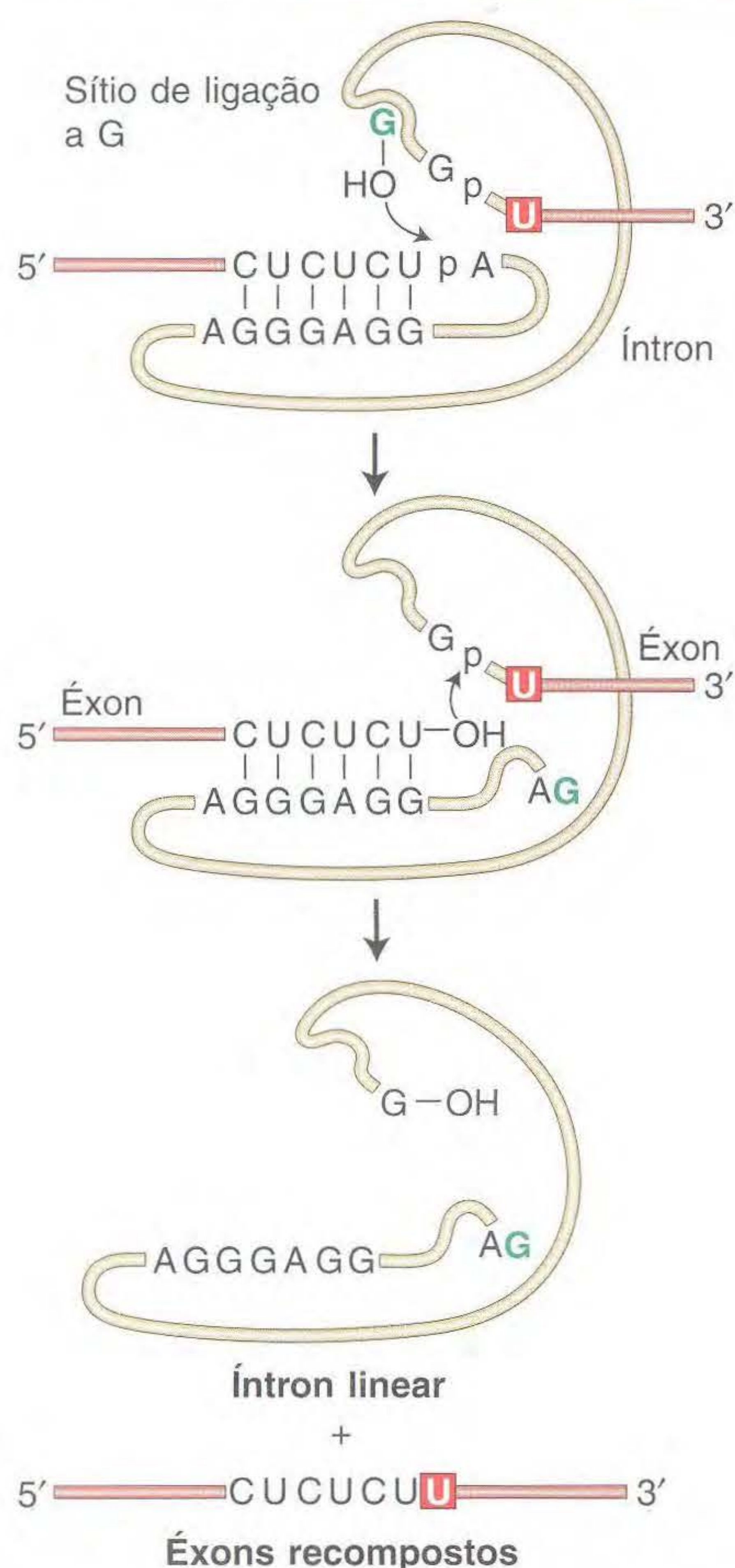


FIG. 8.18 O íntron auto-removível de *Tetrahymena* executa duas reações de transesterificação para se remover do RNA.

A remoção de íntrons e a união de éxons são catalisadas por moléculas de RNA. Em eucariontes, os snRNA do spliceossomo catalisam a remoção dos íntrons do pré-mRNA. Alguns íntrons são auto-removíveis; nesses casos, o próprio íntron catalisa sua remoção. Os RNA capazes de catálise são chamados de ribozimas.

Pequenos RNA de interferência (siRNA)

◆ O QUE OS GENETICISTAS ESTÃO FAZENDO HOJE

Em 2002, uma das principais publicações científicas, a revista *Science*, chamou o “*Small RNA*” como a Conquista do Ano (Figura 8.19). Os RNA aludidos não eram os previamente descritos pequenos RNA, tais como snRNA ou tRNA. Esses RNA foram primeiro descritos nos anos 1990 como parte de duas áreas muito diferentes de experimentação, uma em animais e outra em plantas.

Em 1998, Andrew Fire e Craig Mello relataram que haviam encontrado um potente modo para desligar seletivamente genes no nematelminto *C. elegans*. Fire e Mello descobriram que, injetando cópias de RNA bifilamentar (dsRNA) de um gene de *C. elegans* em embriões de *C. elegans*, eles foram capazes de bloquear a síntese do produto proteico desse gene. O desligamento seletivo do gene por esse procedimento é chamado de **silenciamento gênico** (o silenciamento gênico também pode ocorrer naturalmente; veja o Capítulo 11). O dsRNA tinha sido sintetizado no laboratório e foi composto de um filamento com sentido (codificante) e um filamento de RNA **anti-sentido** complementar. Em seu experimento inicial, Fire e Mello injetaram cópias de dsRNA do gene *unc-22* em embriões de *C. elegans* e observaram que os embriões cresceram em adultos que se contorciam e tinham

ficou uma ribonucleoproteína (chamada RNase P) responsável por cortar a molécula de pré-tRNA em um sítio específico. A grande surpresa veio ao determinar que a atividade catalítica da RNase P residia no componente RNA da enzima, e não no componente proteico. Os achados de Cech e Altman são considerados descobertas marcantes em biologia porque marcaram a primeira vez em que moléculas biológicas que não proteínas eram mostradas catalisando reações.

Essa descoberta de íntrons de auto-replicação levou ao reexame do papel dos snRNA no spliceossomo. Os estudos mais recentes indicam que a remoção do íntron é catalisada pelos snRNA, e não pelo componente proteico do spliceossomo. Como você verá no Capítulo 9, os RNA no ribossomo (os rRNA), não as proteínas ribossômicas, são tidas como tendo um papel central na maioria dos eventos importantes da síntese de proteínas. Os vários exemplos de ribozimas deram uma sólida evidência de uma teoria chamada de o **mundo de RNA**, que diz que o RNA deve ter sido o material genético nas primeiras células porque só o RNA é conhecido como codificando tanto a informação genética quanto catalisando reações biológicas.



FIG. 8.19 Capa da revista *Science* de dezembro de 2002. [De *Science*, vol. 298, no. 5602, December 20, 2002. Reproduzida com permissão da AAAS.]

defeitos musculares. Esse resultado foi entusiasmante porque *unc-22* era conhecido como codificante de uma proteína muscular, e os mutantes nulos de *unc-22* apresentavam a mesma contorção e defeitos musculares. Consideradas juntas, essas observações indicaram que o dsRNA injetado impedia a produção da proteína *Unc-22*. A pergunta que restou era como?

Logo depois, David Baulcombe estava pesquisando o motivo pelo qual as plantas de tabaco que foram modificadas para expressar o gene viral eram resistentes à subsequente infecção pelo vírus. Sabia-se que o vírus não podia replicar seu genoma de RNA nas plantas modificadas. Baulcombe *et al.* descobriram que as plantas resistentes, e apenas as plantas resistentes, produziam grandes quantidades de pequenos RNA, com 25 nucleotídeos de tamanho, que eram complementares ao genoma viral. Os pequenos RNA também eram produzidos durante o silenciamento gênico em *C. elegans*.

A produção de pequenos RNA também foi vista associada a dois fenômenos correlatos em plantas: o **silenciamento transgênico** e a **co-supressão**. O **transgene** é pequeno para o “gene transformado”; um **transgene** é um gene que foi introduzido nos cromossomos de um organismo em laboratório. Às vezes, o transgene é incorporado aos cromossomos, mas não faz um produto proteico. Nesse caso, o silenciamento do transgene é dito como tendo ocorrido. Se o transgene é uma cópia de um gene que é um gene normal no organismo (chamado de um **gene endógeno**), tanto o transgene quanto a cópia normal podem falhar na produção de um produto proteico. Esse resultado é chamado de co-supressão. Por exemplo, os padrões florais incomuns mostrados na Figura 8.20 resultam de co-supressão. Os lindos padrões foram primeiro observados quando o cientista de plantas Richard Jorgenson inseriu um gene de petúnia que codifica uma enzima necessária para a síntese do pigmento floral azul-púrpura em uma petúnia normal (flores azul-púrpura). Dito de outro modo, nesse experimento um transgene foi inserido nos cromossomos de uma petúnia que já tinha uma boa cópia desse gene

em seu locus usual no genoma de petúnia. Em um resultado totalmente inesperado, o transgene de petúnia disparou a supressão tanto do transgene quanto do gene de pigmento normal. Significativamente, os pequenos RNA também tinham sido encontrados presentes durante o silenciamento transgênico e co-supressão.

Em resumo, os pequenos RNA foram observados no silenciamento gênico por dsRNA em *C. elegans*, na resistência viral em plantas de tabaco geneticamente modificadas e no silenciamento transgênico e co-supressão em várias outras espécies de plantas. A produção de pequenos RNA em todos os três fenômenos sugere uma conexão entre esses fenômenos diversos nos reinos vegetal e animal.

A identificação da maquinaria celular que leva ao silenciamento gênico e transgênico e à resistência viral foi uma das áreas mais incrementadas da pesquisa genética na última década. Hoje temos uma melhor compreensão de como dsRNA e os pequenos RNA entram no esquema. Os curtos RNA (21-25 nucleotídeos de tamanho) hoje são coletivamente chamados de *pequenos RNA de interferência*, e o fenômeno que resulta no silenciamento do gene e do transgene, bem como na resistência viral pela produção de siRNA, é chamado de **RNA de interferência (RNAi)**. A Figura 8.21 resume nosso conhecimento atual do que é conhecido como a **via de RNAi**. Resumidamente, duas máquinas biológicas, ambas com a habilidade de quebrar o RNA, tomam parte em um processo de duas etapas. Uma máquina, chamada Dicer, reconhece longas moléculas de dsRNA e as corta em siRNA bifilamentar. Uma segunda máquina, chamada RISC (complexo de silenciamento induzido por RNA), liga-se a esses siRNA e os desenrola em siRNA unifilamentar, de modo que o filamento anti-sentido, ainda ligado a RISC, pode se hibridizar com os mRNA celulares que são complementares. Desse modo, RISC é direcionado para destruir mRNA celulares específicos, o que faz cortando-os em pequenos pedaços.

Vimos que os dsRNA disparam um mecanismo que resulta na destruição de mRNA complementares na célula. No

Flores de petúnia demonstrando co-supressão



(a)



(b)



(c)

FIG. 8.20 (a) O fenótipo selvagem (sem transgene). (b e c) Fenótipos chamados de co-supressão resultantes da transformação da petúnia tipo selvagem mostrada na parte a com um gene de petúnia necessário para pigmentação. Nas regiões sem cor, ambos os transgenes e a cópia cromossômica do mesmo gene foram inativados. [Fotos por cortesia de Richard A. Jorgensen, Department of Plant Sciences, University of Arizona.]

Mecanismo de ação do RNAi

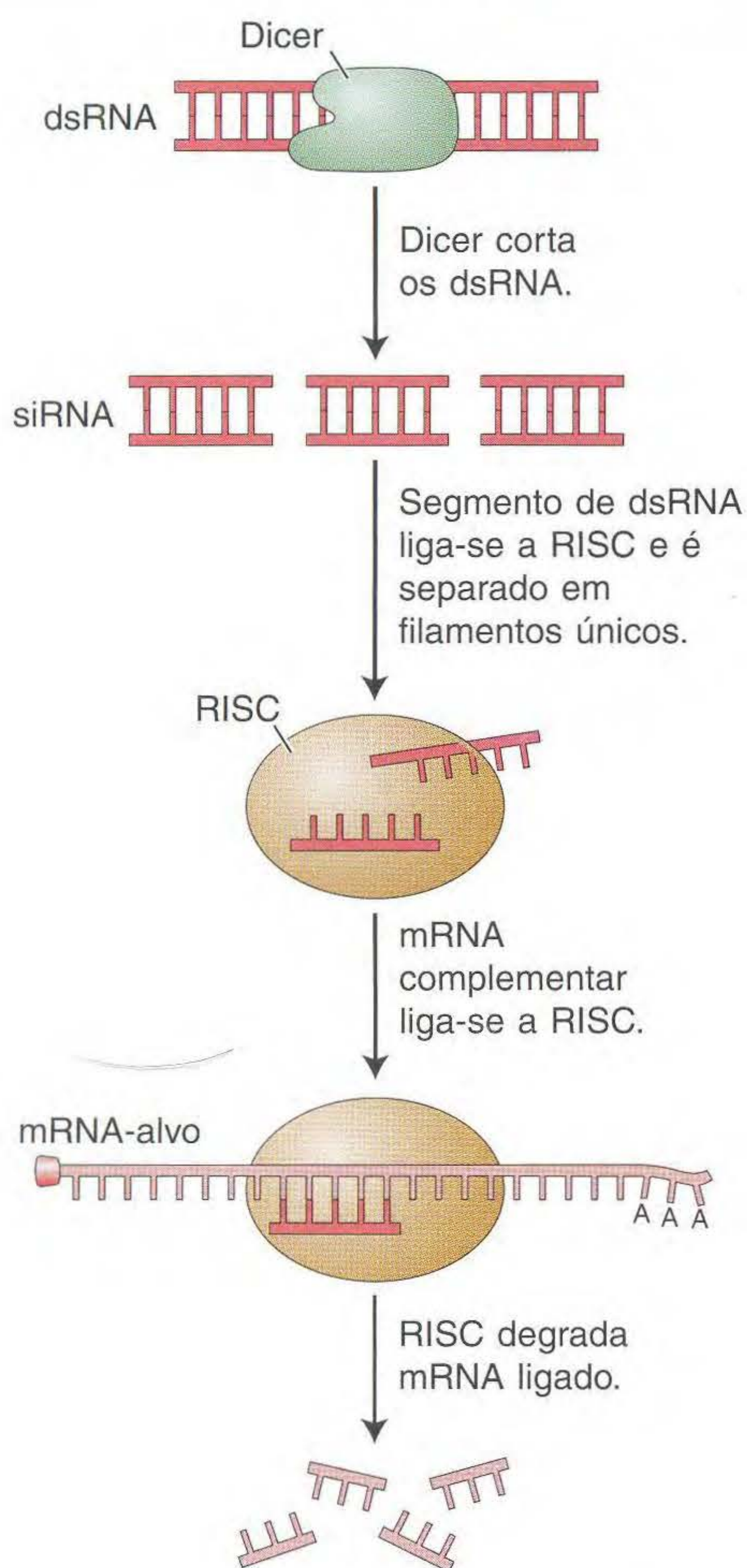


FIG. 8.21 Na via de interferência de RNA, o RNA bifilamentar (dsRNA) interage com o complexo Dicer, que corta o dsRNA. O complexo de silenciamento (RISC) induzido por RNA usa os pequenos dsRNA para encontrar e destruir o mRNA homólogo transcrito do DNA-alvo.

experimento no qual os genes do verme foram silenciados, os dsRNA foram injetados em embriões de laboratório. Quais foram as fontes de dsRNA nas plantas transgênicas que eram resistentes ao vírus ou que sofriam silenciamento transgênico ou co-supressão? Em todos esses casos, as plantas transgênicas contendo genes foram introduzidas no laboratório. Como os cientistas não podem controlar onde os transgenes se inserem, alguns transgenes acabarão perto de genes de plantas em um sentido oposto. A Figura 8.22 ilustra como a co-supressão resulta quando isso ocorre. A transcrição iniciada no promotor do gene pode ser “lida” no transgene e produzir um RNA “quimérico” muito longo contendo tanto

Um modo para gerar RNA bifilamentar em uma célula

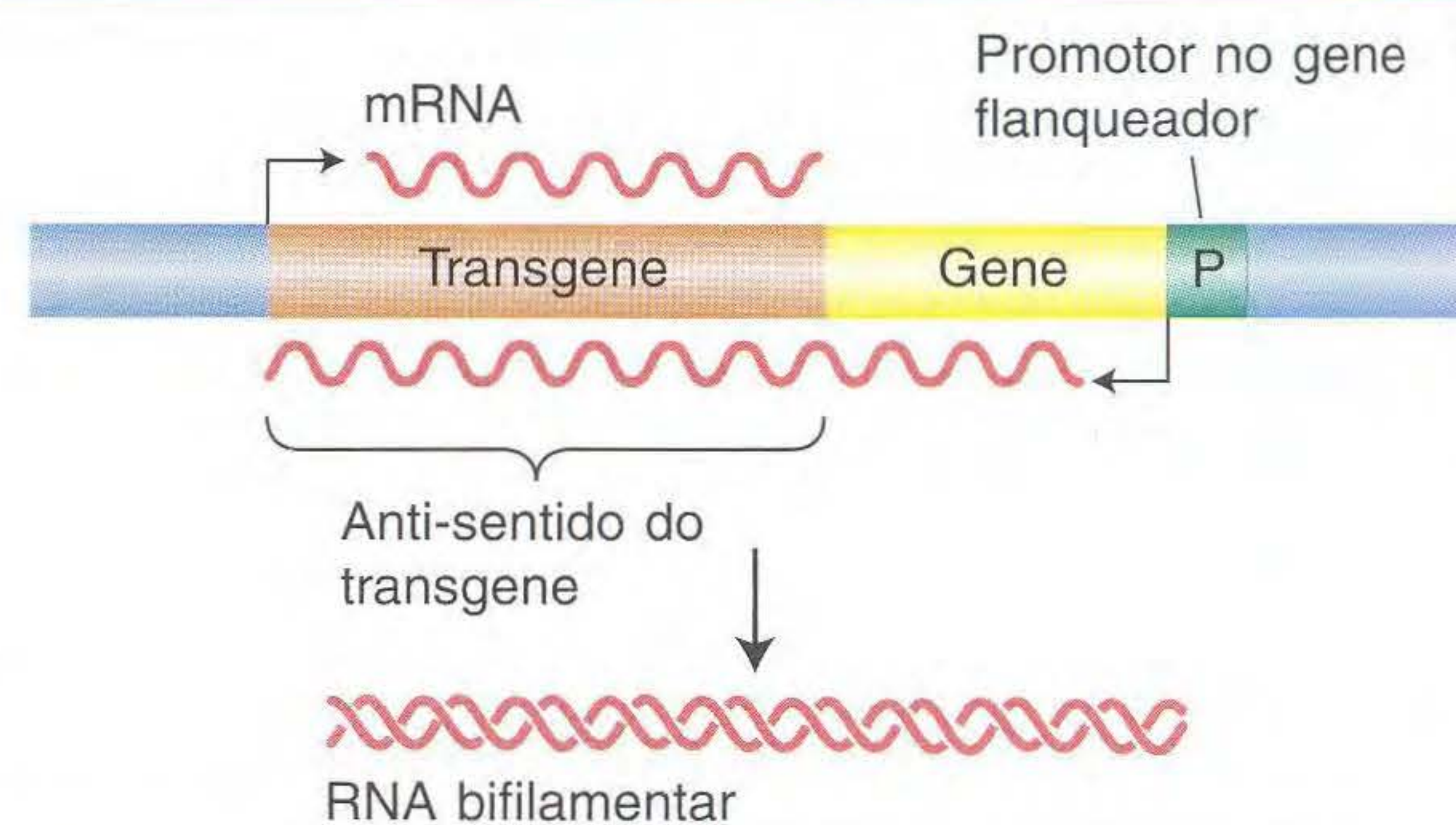


FIG. 8.22 Como descrito no texto, o dsRNA produzido pela injeção direta ou subsequente à integração ao cromossomo então estimulará a via de RNAi.

o filamento com sentido do gene e o filamento anti-sentido do transgene. O RNA bifilamentar será então formado quando a parte anti-sentido do longo RNA se hibridizar com o transgene de RNA com sentido. Dicer reconhece o dsRNA e o corta em siRNA, que irão ligar-se a RISC e dirigir esse complexo para um mRNA complementar na célula. Nos casos aqui descritos, o RNA complementar poderia ser de um vírus infectante (nas plantas resistentes), do próprio transgene (no silenciamento do transgene) ou de um gene endógeno (na co-supressão).

O fato de a via de RNAi ser conservada tanto em plantas quanto em animais indica que é um mecanismo antigo que evoluiu em seu ancestral comum há cerca de um bilhão de anos. Quando um processo biológico é altamente conservado, ele é considerado como fazendo algo muito importante. Quais são os papéis vitais desempenhados pela via de RNAi em condições naturais? Como muitas outras perguntas em genética, essa questão está sob investigação em muitos, talvez em centenas de laboratórios pelo mundo. A via de RNAi provavelmente tem um papel importante na defesa viral. Entretanto, seu principal papel deve ser proteger o material genético de um organismo dos elementos genéticos em seu próprio genoma. No Capítulo 14, você aprenderá sobre os elementos de transposição que constituem uma enorme fração dos genomas de eucariontes multicelulares, inclusive humanos. Esses elementos podem se amplificar e mover-se para novos locais, criando uma ameaça óbvia à integridade do genoma. Do mesmo modo que a introdução dos transgenes pelos cientistas, o movimento dos elementos de transposição para novos locais cromossômicos podem disparar a via de RNAi gerando dsRNA. A via de RNAi ocasionalmente inativa o elemento de transposição e pára a amplificação.

Resumo

Sabemos que a informação não é transferida diretamente do DNA para a proteína, pois, em uma célula eucariótica, o DNA está no núcleo, enquanto a proteína é sintetizada

no citoplasma. A transferência da informação do DNA para uma proteína requer um intermediário. Esse intermediário é o RNA.

Embora o DNA e o RNA sejam ácidos nucleicos, o RNA difere do DNA, (1) pois geralmente é unifilamentar e não uma dupla hélice; (2) seus nucleotídeos contêm o açúcar ribose em vez de desoxirribose; (3) ele tem a base pirimidínica uracil em vez de timina; e (4) ele pode servir como um catalisador biológico.

A similaridade do RNA com o DNA sugere que o fluxo de informação do DNA para o RNA é baseado na complementariedade das bases, o que também é a chave para a replicação do DNA. Um molde de filamento de DNA é copiado, ou transcrito, em um RNA funcional (tal como o RNA transportador ou RNA ribossômico), que nunca é traduzido em polipeptídeos, ou um RNA mensageiro, a partir do qual as proteínas são produzidas.

Em procariontes, todas as classes de RNA são transcritas por uma única RNA polimerase. Essa enzima multissubunitária inicia a transcrição ligando-se ao DNA em promotores que contêm seqüências específicas em -35 e -10 bases antes do início da transcrição em +1. Após ser ligada, a RNA polimerase desenrola localmente o DNA e começa a incorporar ribonucleotídeos que são complementares ao filamento molde de DNA. O crescimento da cadeia no sentido 5' para 3' faz-se até que um dos dois mecanismos, intrínseco ou dependente de *rô*, leve à dissociação da polimerase e RNA do molde de DNA. Como veremos no Capítulo 9, na ausência de um núcleo, os RNA procarióticos que codificam proteínas são traduzidos enquanto estão sendo transcritos.

Em eucariontes, existem três RNA polimerases diferentes; apenas a RNA polimerase II transcreve mRNA. No geral, as fases de iniciação, alongamento e término da síntese de RNA em eucariontes lembram aquelas em procariontes. Entre-

tanto, existem diferenças importantes. A RNA polimerase II não se liga diretamente ao DNA promotor, mas sim a fatores gerais de transcrição, um dos quais reconhece a seqüência TATA na maioria dos promotores eucarióticos. A RNA polimerase II é uma molécula muito maior que sua contraparte procariótica. Ela contém várias subunidades que funcionam não só para alongar o RNA transcrito primário, mas também para coordenar os amplos eventos de processamento que são necessários para produzir o mRNA final. Esses eventos de processamento são revestimento em 5', remoção de íntrons e união de éxons pelos spliceossomos, e clivagem em 3' seguida de poliadenilação. Parte do cerne da RNA polimerase II, o domínio da cauda carboxílica, é idealmente posicionado para interagir com o RNA nascente à medida que ele emerge da polimerase. Pelo CTD, a RNA polimerase II coordena os vários eventos de síntese e processamento do RNA.

As descobertas dos últimos 20 anos revelaram a importância das novas classes de RNA funcional. Antes imaginado como um mensageiro solitário, o RNA hoje é reconhecido como um participante dinâmico e versátil em muitos processos celulares. A descoberta dos íntrons auto-recompostos demonstrou que o RNA pode funcionar como um catalisador, como as proteínas. Desde a descoberta dessas ribozimas, a comunidade científica começou a dar mais atenção ao RNA. Os pequenos RNA nucleares, os RNA não-codificantes no spliceossomo, hoje são reconhecidos como dando a atividade catalítica para remover os íntrons e unir os éxons. O século XX terminou com a descoberta de que outra classe de RNA funcional, o pequeno RNA de interferência, protege a integridade dos genomas de plantas e animais ativando a via de interferência do RNA.

Termos-chave

alongamento	molde	RNA de interferência (RNAi)
antecedente	mundo de RNA	RNA de transferência (tRNA)
bolha de transcrição	pequeno RNA de interferência (siRNA)	RNA funcional
<i>cap</i>	pequeno RNA nuclear (snRNA)	RNA mensageiro (mRNA)
cauda poli(A)	posterior	RNA polimerase
complexo de pré-iniciação (PIC)	processamento co-transcricional	RNA polimerase holoenzima
co-supressão	processamento de RNA	RNA ribossômico (rRNA)
domínio da cauda carboxílica (CTD)	processamento pós-transcricional	seqüência de consenso
éxon	promotor	silenciamento gênico
experimento de pulso-caça	proteína de ligação a TATA (TBP)	silenciamento transgênico
fator geral de transcrição (GTF)	proteoma	spliceossomo
fator sigma (σ)	recomposição	TATA boxe
filamento codificante	recomposição alternativa	término
filamento de RNA anti-sentido	recomposição de RNA	transcrição
gene endógeno	região não-traduzida 3' (3' UTR)	transcrito
iniciação	região não-traduzida 5' (5' UTR)	transcrito primário (pré-mRNA)
íntron	regra GU-AG	transgene
íntron de auto-remoção	ribose	uracil (U)
microRNA (miRNA)	ribozima	via de RNAi

► Problemas

PROBLEMAS BÁSICOS

- Os dois filamentos do DNA do fago λ diferem um do outro em seu conteúdo de GC. Devido a essa propriedade, eles podem ser separados em um gradiente alcalino de cloreto de cério (a alcalinidade desnatura a dupla hélice). Quando o RNA sintetizado pelo fago λ é isolado de células infectadas, ele é encontrado formando híbridos DNA-RNA com ambos os filamentos do DNA λ . O que esse achado lhe diz? Formule algumas previsões testáveis.
- Tanto em procariontes quanto em eucariontes, descreva o que mais está ocorrendo ao RNA enquanto a RNA polimerase está sintetizando um transcrito do molde de DNA.
- Cite três exemplos de proteínas que atuam nos ácidos nucleicos.
- Qual a função primária do fator sigma? Existe uma proteína nos eucariontes análoga ao fator sigma?
- Você identificou uma mutação em levedura, um eucarionte unicelular, que impede o revestimento da ponta 5' do RNA transcrito. Entretanto, para sua surpresa, todas as enzimas necessárias para o revestimento são normais. Você determinou que a mutação é em uma das subunidades da RNA polimerase II. Que subunidade é mutante e como essa mutação resulta na falha de adição de um revestimento no RNA de levedura?
- Por que o RNA é produzido apenas a partir do DNA-molde e não de ambos os filamentos?
- Um plasmídeo linear contém apenas dois genes, que são transcritos em sentidos opostos, cada um a partir de uma ponta, para o centro do plasmídeo. Desenhe diagramas de
 - DNA do plasmídeo, mostrando as pontas 5' e 3' dos filamentos de nucleotídeos.
 - o filamento-molde para cada gene.
 - as posições do sítio de início da transcrição.
 - os transcritos, mostrando as pontas 5' e 3'.
- Existem similaridades entre as bolhas de replicação de DNA e as bolhas de transcrição encontradas em eucariontes? Explique.
- Quais das seguintes frases são verdadeiras sobre o mRNA eucariótico?
 - O fator sigma é essencial para o início correto da transcrição.
 - O processamento do mRNA nascente pode começar antes que sua transcrição esteja completa.
 - O processamento ocorre no citoplasma.
 - O término é feito pelo uso de um grampo ou pelo uso do fator ρ .
 - Muitos RNA podem ser transcritos simultaneamente a partir de um molde de DNA.

- Uma pesquisadora estava mutando células procarióticas inserindo segmentos de DNA. Desse modo, ela fez as seguintes mutações:

Original	TTGACAT <u>15 a 17 pb</u> TATAAT
Mutante	TATAAT <u>15 a 17 pb</u> TTGACAT

- O que essa sequência representa?
 - Qual efeito você prevê para tal mutação? Explique.
- Você aprenderá mais sobre a engenharia genética no Capítulo 20, mas, por enquanto, seja um engenheiro genético e tente resolver este problema. *E. coli* é amplamente usada nos laboratórios para produzir proteínas a partir de outros organismos.
 - Você isolou um gene de levedura que codifica uma enzima metabólica e quer produzir essa enzima em *E. coli*. Você suspeita que o promotor de levedura não funcionará em *E. coli*. Por quê?
 - Após substituir o promotor de levedura por um promotor de *E. coli*, você ficou satisfeito em detectar RNA do gene de levedura, mas confuso porque o RNA tinha quase o dobro do tamanho do mRNA desse gene isolado de levedura. Explique por que ocorreu esse resultado.
 - Desenhe um gene procariótico e seu produto de RNA. Esteja certo em incluir o promotor, o sítio de início da transcrição, o sítio de término da transcrição, as regiões não-traduzidas e as pontas 5' e 3' marcadas.
 - Desenhe um gene eucariótico com dois íntrons e seu pré-mRNA e mRNA produzido. Esteja seguro de incluir todas as características do gene procariótico incluído em sua resposta ao Problema 12, mais os eventos processantes necessários para produzir o mRNA.
 - Determinado gene codificante de proteína de *Drosophila* tem um íntron. Se uma grande amostra de alelos nulos desse gene é examinada, alguns sítios mutantes serão esperados.
 - nos éxons?
 - no íntron?
 - no promotor?
 - no limite íntron-éxon?
 - O que são íntrons auto-recompostos e por que a sua existência apóia a teoria de que o RNA surgiu antes das proteínas?
 - Antibióticos são drogas que matam seletivamente as bactérias sem prejudicar os animais. Muitos antibióticos atuam seletivamente em determinadas proteínas que são críticas para o funcionamento da bactéria. Explique por que alguns dos mais bem-sucedidos antibióticos têm como o alvo a RNA polimerase.

PROBLEMAS DESAFIADORES

- Os seguintes dados representam as composições básicas do DNA bifilamentar de duas espécies bacterianas diferentes e seus produtos de RNA obtidos em experimentos conduzidos *in vitro*:

Espécie	(A + T)/ (G + C)	(A + U)/ (G + C)	(A + G)/ (U + C)
<i>Bacillus subtilis</i>	1,36	1,30	1,02
<i>E. coli</i>	1,00	0,98	0,80

a. A partir desses dados, você pode determinar se o RNA dessa espécie é copiado de um único filamento ou de ambos os filamentos do DNA? Como? Desenhar um diagrama facilitará resolver esse problema.

b. Explique como você pode dizer se o próprio RNA é unifilamentar ou bifilamentar.

(O Problema 17 é reimpresso com a permissão de Macmillan Publishing Co., Inc. de M. Strickberger, *Genetics*. Copyright 1968, Monroe W. Strickberger.)

18. Um gene humano foi inicialmente identificado como tendo três éxons e dois íntrons. Os éxons têm 456, 224 e 524 pb, enquanto os íntrons têm 2,3 kb e 4,6 kb.

a. Desenhe esse gene, mostrando o promotor, os íntrons, éxons e os sítios de início e fim da transcrição.

b. Surpreendentemente, esse gene foi encontrado codificando não um mas dois mRNA que têm apenas 224 nucleotídeos em comum. O mRNA original tem 1.204 nucleotídeos, e o novo mRNA tem 2.524 nucleotídeos. Use seu desenho para mostrar como essa região do DNA codifica esses dois transcritos.

19. Enquanto está trabalhando em seu laboratório, você isola um mRNA de *C. elegans* que suspeita ser essencial para os embriões se desenvolverem bem-sucedidamente. Com a suposição de que você é capaz de transformar um mRNA em um RNA bifilamentar, crie um experimento para testar sua hipótese.

20. Glifosato é um herbicida usado para matar pragas. É o principal componente de um produto feito pela Monsanto Company chamado Roundup. O glifosato mata as plantas inibindo uma enzima na via shikimate chamada

EPSPS. Esse herbicida é considerado seguro porque os animais não têm a via shikimate. Para vender ainda mais de seu herbicida, a Monsanto mandou seus geneticistas de plantas fazerem vários cultivos, inclusive de milho, para serem resistentes ao glifosato. Para fazer isso, os cientistas introduziram a enzima EPSPS, que era resistente à inibição pelo glifosato, em um cultivo de plantas e, então, testaram as plantas transformadas para resistência ao herbicida.

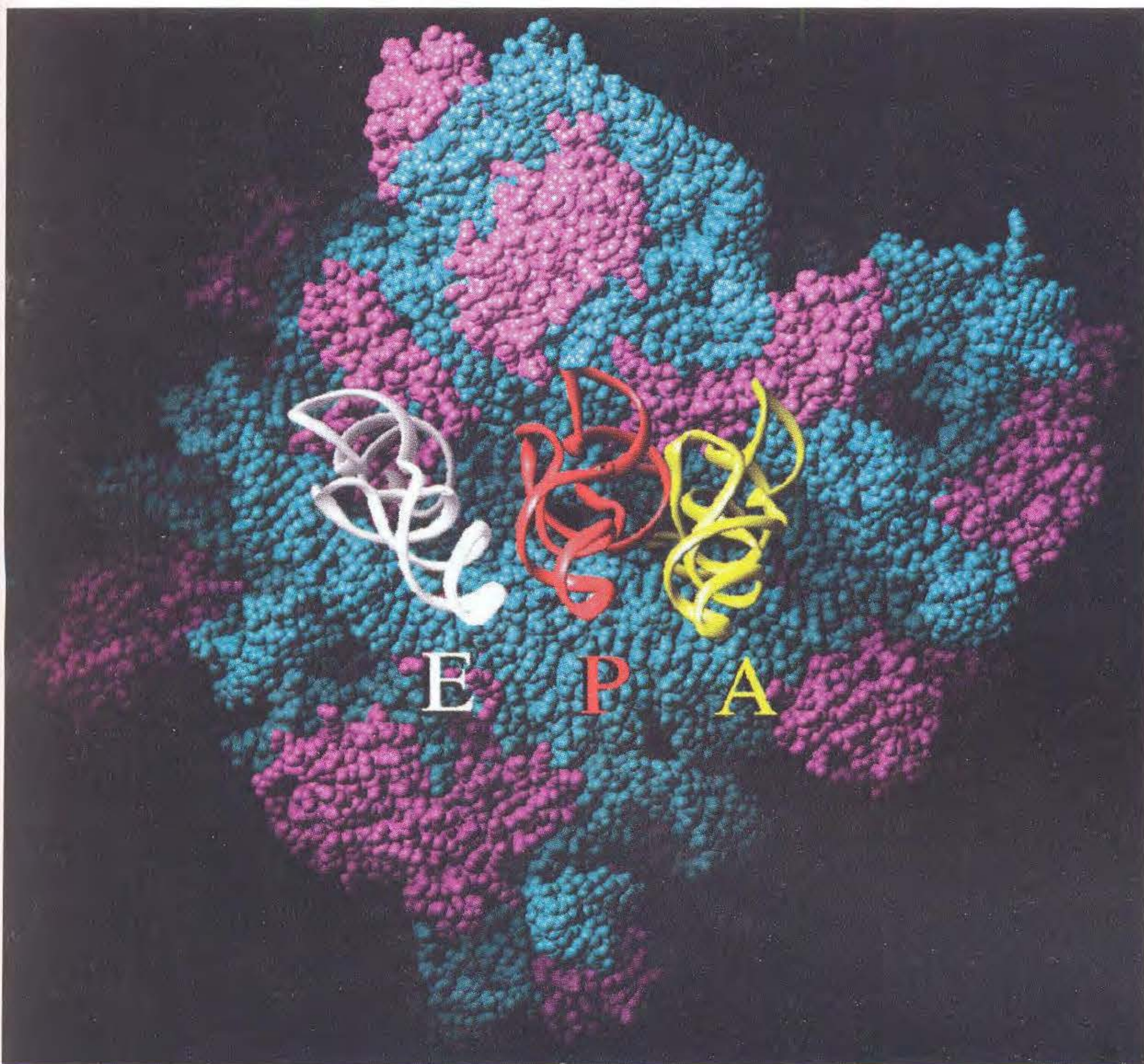
Imagine que você é um desses cientistas e que teve sucesso em introduzir o gene resistente EPSPS nos cromossomos de milho. Você descobre que algumas das plantas transgênicas são resistentes ao herbicida, enquanto outras não. Seu supervisor não gostou e pede uma explicação do motivo pelo qual algumas plantas não são resistentes muito embora tenham o transgene em seus cromossomos. Desenhe uma figura para ajudá-lo a entender.

21. Muitos cânceres humanos resultam de uma mutação em um gene normal que leva a um crescimento desordenado (um tumor). Os genes que causam câncer quando mutados são chamados de oncogenes. A quimioterapia é efetiva contra muitos tumores porque tem como alvo células que se dividem rapidamente e as matam. Infelizmente, a quimioterapia tem muitos efeitos colaterais, tais como a perda dos cabelos ou náuseas, porque também mata muitas de nossas células normais que estão se dividindo rapidamente, tais como as dos folículos pilosos ou revestimento do estômago.

Muitos cientistas e grandes empresas farmacêuticas estão entusiasmados quanto às perspectivas de explorar a via do RNAi para inibir seletivamente os oncogenes nos tumores que ameaçam a vida. Explique em termos bem gerais como a terapia de silenciamento gênico pode funcionar para tratar o câncer, e por que esse tipo de terapia teria menos efeitos colaterais que a quimioterapia.

Proteínas e sua Síntese

9



Esta imagem mostra em resolução atômica uma superfície do ribossomo da bactéria *Halobacterium marismortui*, deduzida da cristalografia de raios X. A parte do ribossomo que consiste em RNA é mostrada em azul; a que consiste em proteína é mostrada em púrpura. As estruturas brancas, vermelhas e amarelas no centro são tRNA nos sítios de ligação E, P e A, e suas hastes aceptoras desaparecem em uma fenda no ribossomo. [De P. Nissen, J. Hansen, N. Ban, P.B. Moore, and T.A. Steitz, "The Structural Basis of Ribosome Activity in Peptide Bond Synthesis," *Science* 289, 2000, 920–930, Fig. 10A na p. 926.]

Perguntas Fundamentais

- Como são relacionadas as seqüências de um gene e suas proteínas correlatas?
- Por que se diz que o código genético é não-superposto e redundante?
- Qual a evidência de que o RNA ribossômico, e não as proteínas ribossômicas, efetua as principais etapas na tradução?
- O que é processamento pós-traducional, e por que ele é importante para o funcionamento da proteína?

Tópicos

- 9.1 Estrutura das proteínas
- 9.2 Colinearidade de gene e proteína
- 9.3 O código genético
- 9.4 tRNA: o adaptador
- 9.5 Ribossomos
- 9.6 O proteoma

Em uma mensagem ao Congresso em 1969, William Stewart, cirurgião geral dos EUA, disse: "É a hora de fechar o livro das doenças infecciosas. A guerra contra a pestilência acabou." Naquela época, sua proclamação de vitória não era descabida. Nas duas décadas precedentes, três doenças infecciosas que tinham assolado a humanidade por séculos, a pólio, varíola e tuberculose, tinham sido praticamente eliminadas do mundo. Um importante fator contribuinte para a erradicação de algumas doenças infecciosas foi a descoberta e amplo uso dos *antibióticos*, um grupo diverso de compostos químicos que matam patógenos bacterianos específicos sem prejudicar o hospedeiro animal. Os antibióticos tais como a penicilina, tetraciclina, ampicilina e cloranfenicol, para citar apenas alguns, salvaram centenas de milhões de vidas.

Infelizmente, a exclamação de vitória de William Stewart na batalha contra doenças infecciosas foi prematura. O uso excessivo de antibióticos em

todo o mundo provocou a evolução de linhagens bacterianas resistentes. Por exemplo, a cada ano, mais de 2 milhões de pacientes de hospitais nos USA contraem uma infecção que é resistente a antibióticos, e 90.000 morrem como resultado. Como a resistência desenvolveu-se tão rapidamente? As doenças infecciosas serão, novamente, uma causa significativa de mortalidade humana? Ou os cientistas serão capazes de usar sua compreensão dos mecanismos de resistência para desenvolver antibióticos mais duráveis?

Para responder a essas perguntas, os cientistas têm focado a maquinaria celular que é o alvo dos antibióticos. Mais de metade de todos os antibióticos atualmente em uso têm como alvo o ribossomo bacteriano, o sítio da síntese de proteínas nos procariontes. Neste capítulo, você aprenderá que os cientistas tiveram um incrível sucesso com o uso de uma técnica chamada cristalografia de raios X para visualizar os RNA ribossômicos (rRNA) e as cerca de 100 proteínas que constituem as subunidades maior e menor do ribossomo bacteriano. Embora os ribossomos dos procariontes e eucariontes sejam muito similares, ainda existem pequenas diferenças. Devido a essas diferenças, os antibióticos são capazes de ter como alvo ribossomos bacterianos, mas deixar intocáveis os ribossomos eucarióticos. Usando a cristalografia de raios X, os cientistas também tiveram sucesso em visualizar antibióticos ligados ao ribossomo (Figura 9.1). Por esses estudos, eles determinaram que as mutações no rRNA bacteriano e/ou proteínas ribossômicas são responsáveis pela resistência

antibiótica. Com esse conhecimento dos pontos de contato entre alguns antibióticos e o ribossomo, os projetistas de drogas estão tentando criar uma nova geração de antibióticos que, por exemplo, sejam capazes de se ligar a vários sítios vizinhos. A lógica de tal projeto de droga é que a resistência seria mais durável porque precisaria da ocorrência de duas mutações, o que é um evento muito improvável, mesmo para uma bactéria.

Os Capítulos 7 e 8 descreveram como o DNA é copiado de geração a geração, e como o RNA é sintetizado a partir de regiões específicas do DNA. Podemos pensar nesses processos como dois estágios da transferência de informação: *replicação* (a síntese de DNA) e a *transcrição* (a síntese de uma cópia de RNA de uma parte do DNA). Neste capítulo, você aprenderá sobre o estágio final da transferência de informação: a *tradução* (a síntese de um polipeptídeo dirigida pela sequência de RNA).

Como você aprendeu no Capítulo 8, o RNA transcrito dos genes é classificado como RNA mensageiro (mRNA) ou RNA funcional. Neste capítulo, veremos o destino de ambas as classes de RNA. A grande maioria dos genes codificam mRNA cuja função é servir como um intermediário na síntese do produto gênico final, a proteína. Em contraste, lembre que os RNA *funcionais* são ativos como RNA; eles nunca são traduzidos em proteínas. As principais classes de RNA funcionais são atores importantes na síntese de proteínas. Elas incluem os RNA transportadores e os RNA ribossômicos.

A ligação de uma molécula de uma substância ao ribossomo impede a tradução

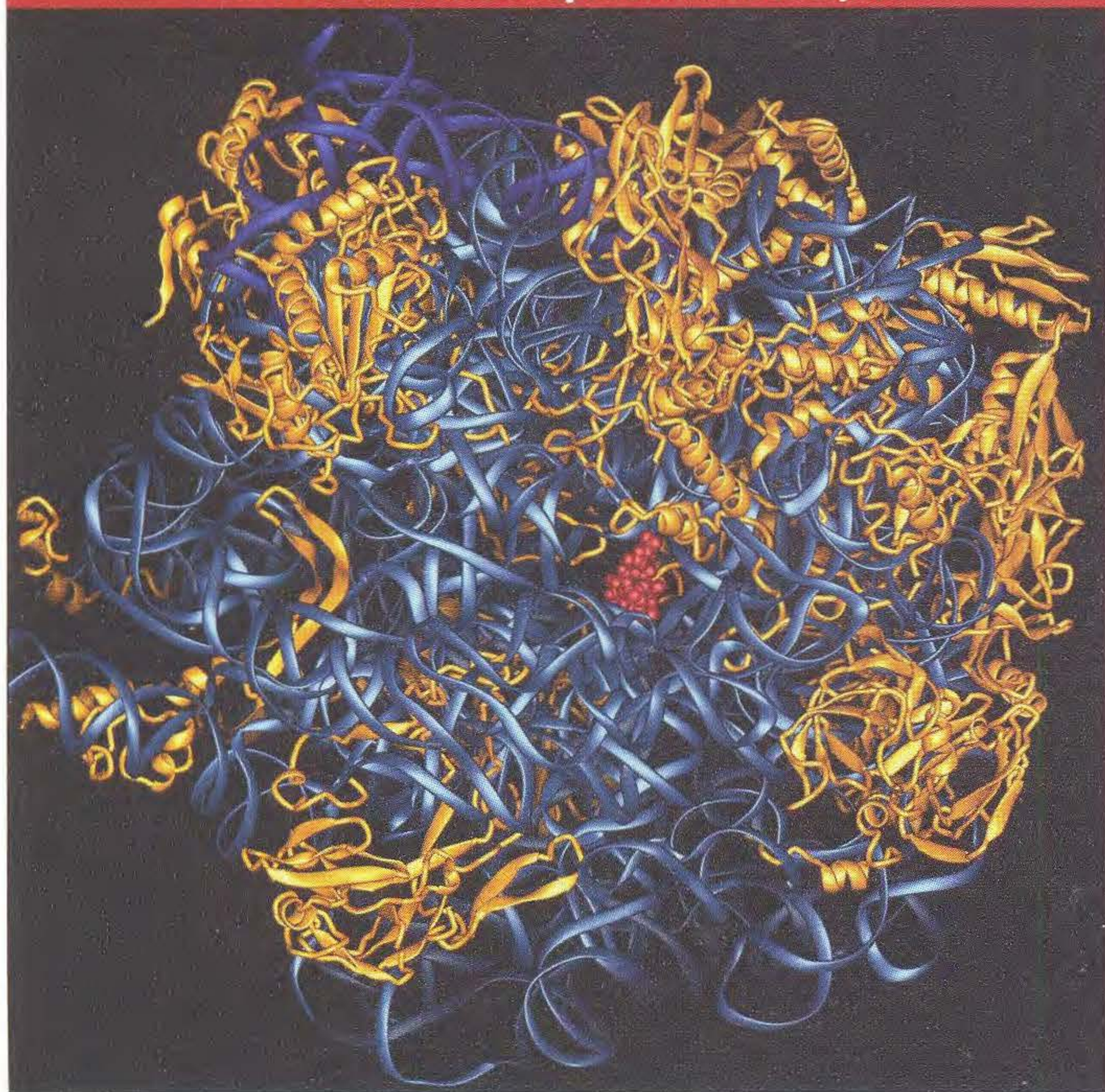


FIG. 9.1 A droga eritromicina (vermelho) bloqueia o túnel através do qual emerge do ribossomo uma proteína recém-sintetizada. A imagem é uma visão de cima da subunidade ribossomal 50S na bactéria *Deinococcus radiodurans*. Os RNA ribossômicos são mostrados em azul, e as proteínas ribossômicas em amarelo. [Dr. Joerg Harms, MPI for Molecular Genetics, Berlin, Germany.]

- As moléculas de **RNA transportador (tRNA)** são os adaptadores do códon de três nucleotídeos no mRNA ao aminoácido correspondente, que é levado pelo tRNA ao ribossomo no processo de tradução. Os tRNA são componentes gerais da maquinaria de tradução; uma molécula de tRNA pode levar um aminoácido ao ribossomo para a finalidade de traduzir *qualquer* mRNA.
- Os **RNA ribossômicos (rRNA)** são os principais componentes dos **ribossomos**, que são grandes complexos macromoleculares que reúnem os aminoácidos para formar a proteína cuja sequência é codificada em um mRNA específico. Os ribossomos são compostos de vários tipos de rRNA e proteínas diferentes. Como o tRNA, os ribossomos têm uma função geral, pois podem ser usados para traduzir os mRNA de *qualquer* gene codificante de proteínas.

Embora a maioria dos genes codifiquem mRNA, os RNA funcionais constituem a maior fração do RNA celular total. Em uma célula eucariótica ativamente em divisão, o rRNA e o tRNA contribuem com quase 95% do RNA total, enquanto o mRNA responde apenas por cerca de 5%. Dois fatores explicam a abundância de rRNA e tRNA. Primeiro, eles são muito mais estáveis do que os mRNA, e, assim, essas moléculas permanecem intactas por mais tempo. Segundo, a transcrição dos genes de rRNA e tRNA constituem mais de metade da transcrição nuclear total em células eucarióticas ativas, e quase 80% da transcrição em células de leveduras.

Os componentes da maquinaria de tradução e o processo de tradução são muito similares em procariontes e eucariontes. A principal característica que distingue a tradução

em procariontes daquela em eucariontes é o local onde a transcrição e a tradução ocorrem na célula: os dois processos ocorrem no mesmo compartimento nos procariontes, enquanto estão fisicamente separados nos eucariontes. Após um amplo processamento, os mRNA eucarióticos são exportados do núcleo para a tradução nos ribossomos que residem no citoplasma. Em contraste, a transcrição e a tradução são acopladas nos procariontes: a tradução de um RNA começa em sua ponta 5' enquanto o resto do mRNA ainda está sendo transcrito.

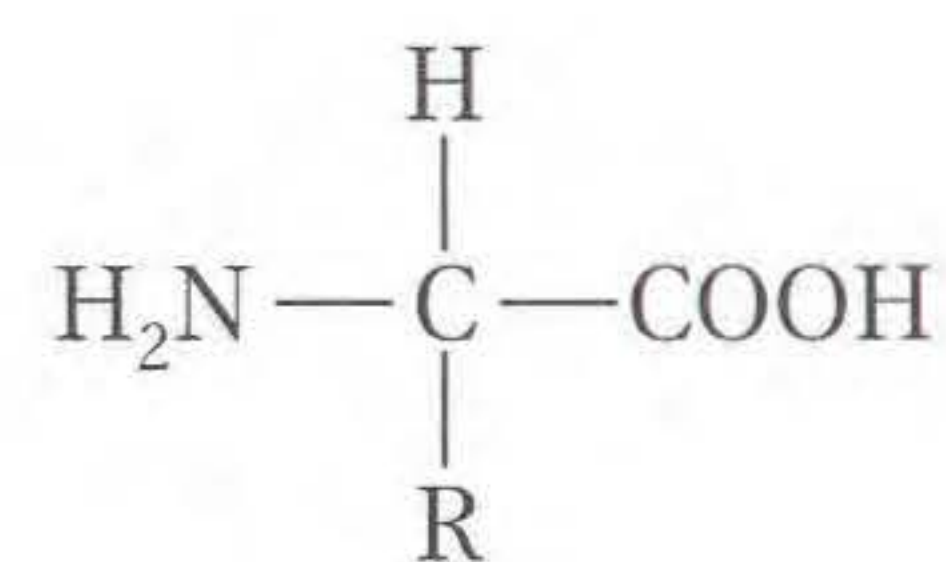
9.1 Estrutura das Proteínas

Quando um transcrito primário foi totalmente processado em uma molécula final de mRNA, pode ocorrer a tradução em proteína. Antes de considerarmos como as proteínas são feitas, precisamos compreender a estrutura da proteína.

As proteínas são os principais determinantes da forma e função biológica. Essas moléculas influenciam muito a

forma, cor, tamanho, comportamento e fisiologia dos organismos. Como os genes funcionam codificando proteínas, a compreensão da natureza das proteínas é essencial para compreender a ação gênica.

Uma proteína é um polímero composto de monômeros chamados de **aminoácidos**. Em outras palavras, uma proteína é uma cadeia de aminoácidos. Como os aminoácidos já foram chamados de *peptídeos*, a cadeia às vezes é chamada de um **polipeptídeo**. Os aminoácidos todos têm uma fórmula geral



Todos os aminoácidos têm uma cadeia lateral, ou grupo R (reativo). Existem 20 aminoácidos conhecidos nas proteínas, cada um tendo um grupo R diferente que dá ao aminoácido suas propriedades únicas. A cadeia lateral pode ser qualquer coisa desde um átomo de hidrogênio (como o aminoácido glicina) até um anel complexo (como o aminoácido triptofano). Nas proteínas, os aminoácidos são unidos por ligações covalentes chamadas ligações peptídicas. Uma ligação peptídica é formada pela união de uma **ponta amino** (NH₂) de um aminoácido com a **ponta carboxila** (COOH) de outro aminoácido. É removida uma molécula de água durante a reação (Figura 9.2). Devido ao modo pelo qual se forma a ligação peptídica, uma cadeia polipeptídica sempre tem uma ponta amino (NH₂) e uma ponta carboxila (COOH), como mostrado na Figura 9.2a.

As proteínas têm uma estrutura complexa com quatro níveis de organização, ilustrados na Figura 9.3. A sequência linear de aminoácidos em uma cadeia polipeptídica constitui uma **estrutura primária** da proteína. As regiões locais da cadeia polipeptídica dobram-se em formas específicas, chamadas de **estrutura secundária** da proteína. Cada forma surge de forças específicas de ligação entre os aminoácidos que estão próximos na sequência linear. Essas forças incluem vários tipos de ligações fracas, notadamente pontes de hidrogênio, forças eletrostáticas e forças de van der Waals. As estruturas secundárias mais comuns são a α hélice e a folha pregueada. Proteínas diferentes apresentam uma ou outra, ou às vezes ambas em suas estruturas. A **estrutura terciária** é produzida pelo dobramento da estrutura secundária. Algumas proteínas têm uma **estrutura quaternária**: uma proteína assim é composta de dois ou mais polipeptídeos separados dobrados, também chamados **subunidades**, unidos por pontes fracas. A associação quaternária pode ser entre tipos diferentes de polipeptídeos (resultando em um heterodímero se existirem duas subunidades) ou entre polipeptídeos idênticos (fazendo um heterodímero). A hemoglobina é exemplo de um heterotetrâmero, uma proteína de quatro subunidades. Ela é composta de duas cópias cada de dois polipeptídeos diferentes, mostrados em verde e púrpura na Figura 9.3d.

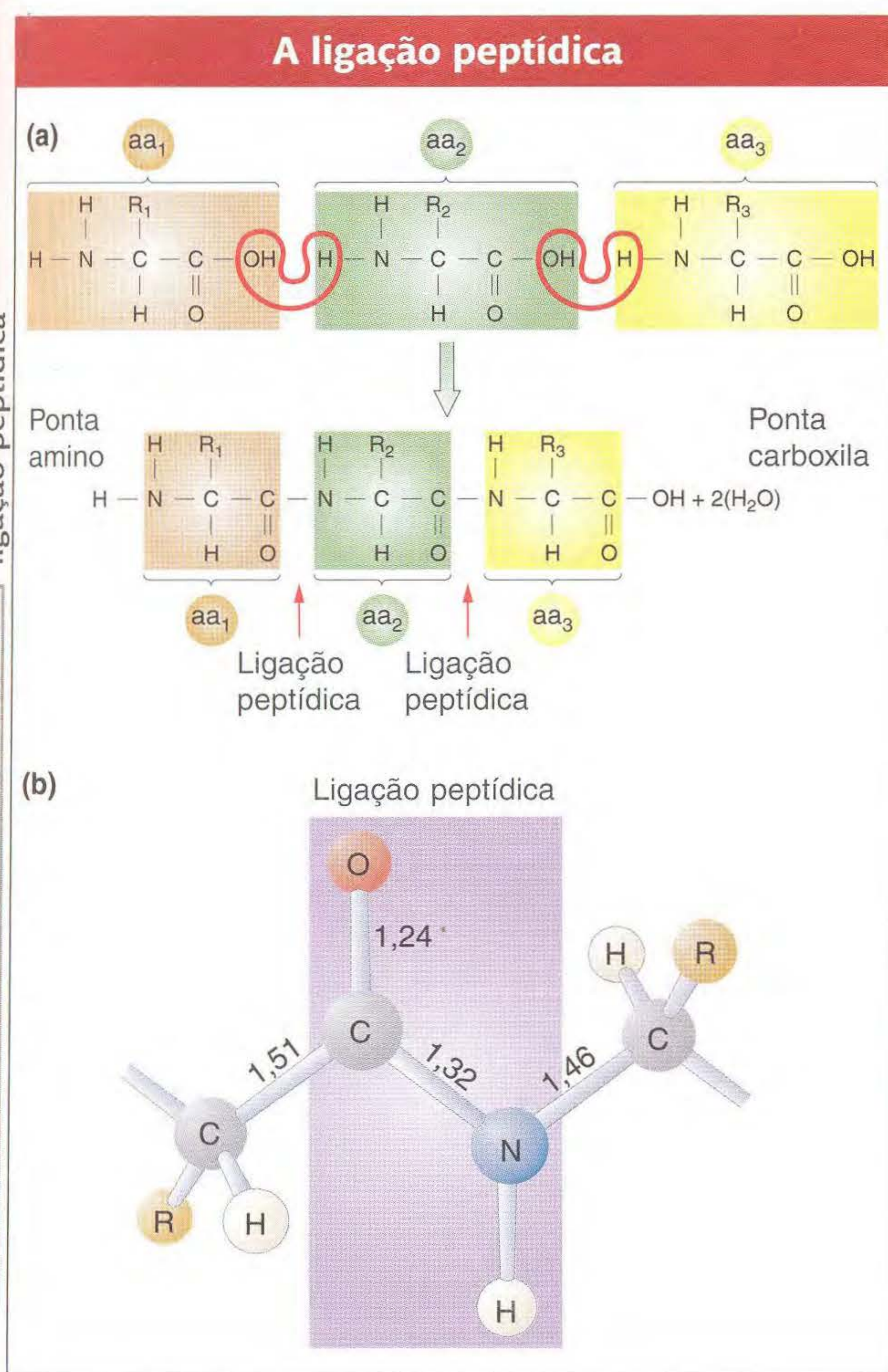
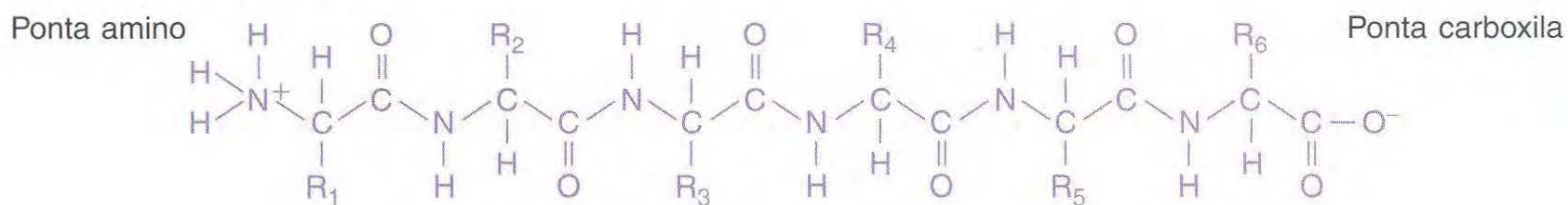


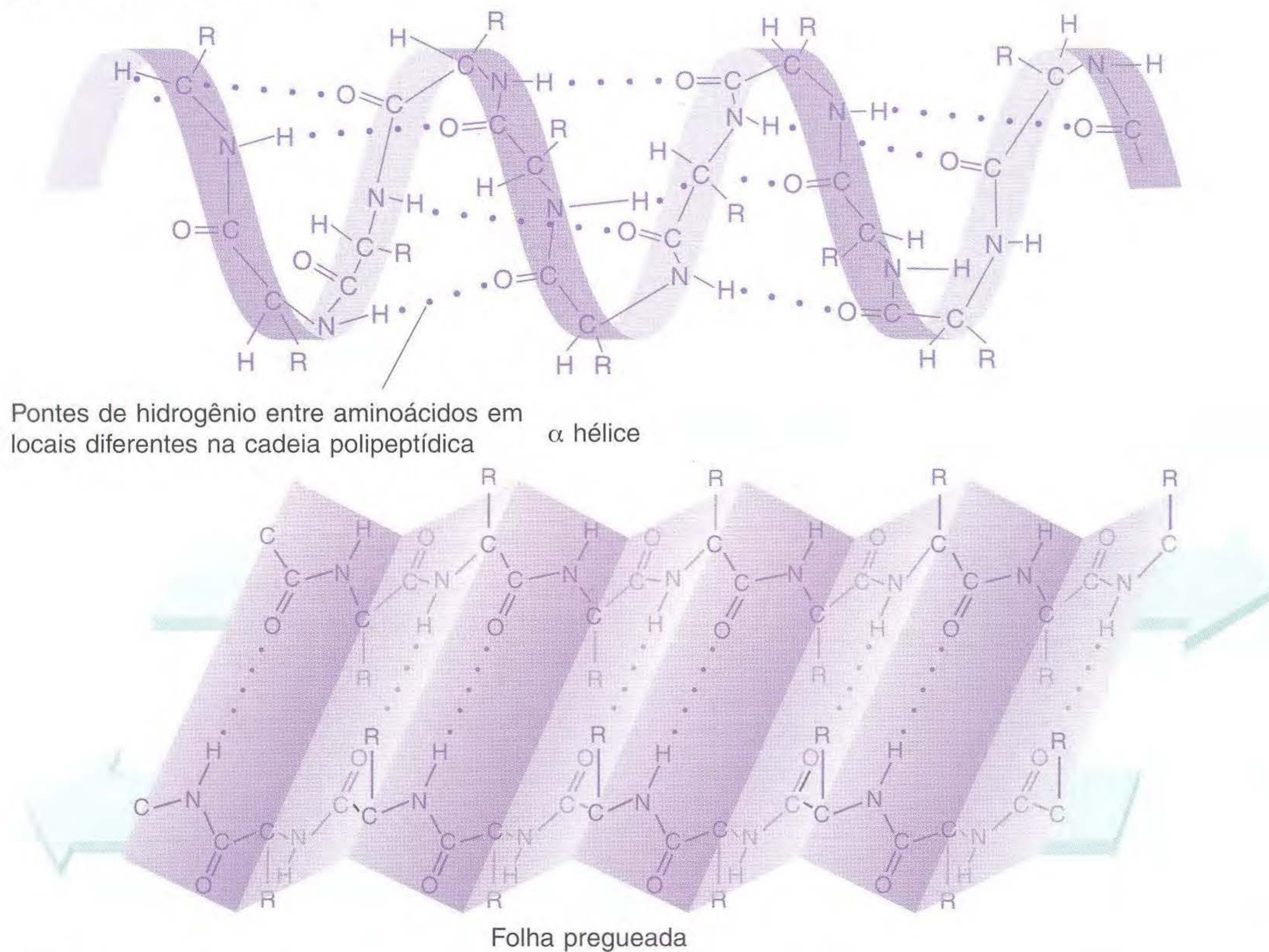
FIG. 9.2 (a) Um polipeptídeo é formado pela remoção de água entre os aminoácidos para formar ligações peptídicas. Cada aa indica um aminoácido. R₁, R₂ e R₃ representam grupos R (cadeias laterais) que diferenciam os aminoácidos. (b) A ligação peptídica é uma unidade planar rígida com os grupos R projetando-se para fora do arcabouço C-N. São mostradas as distâncias padrão das ligações (em ângstrons). [(b) De L. Stryer, *Biochemistry*, 4th ed. Copyright 1995 by Lubert Stryer.]

Níveis de estrutura de proteína

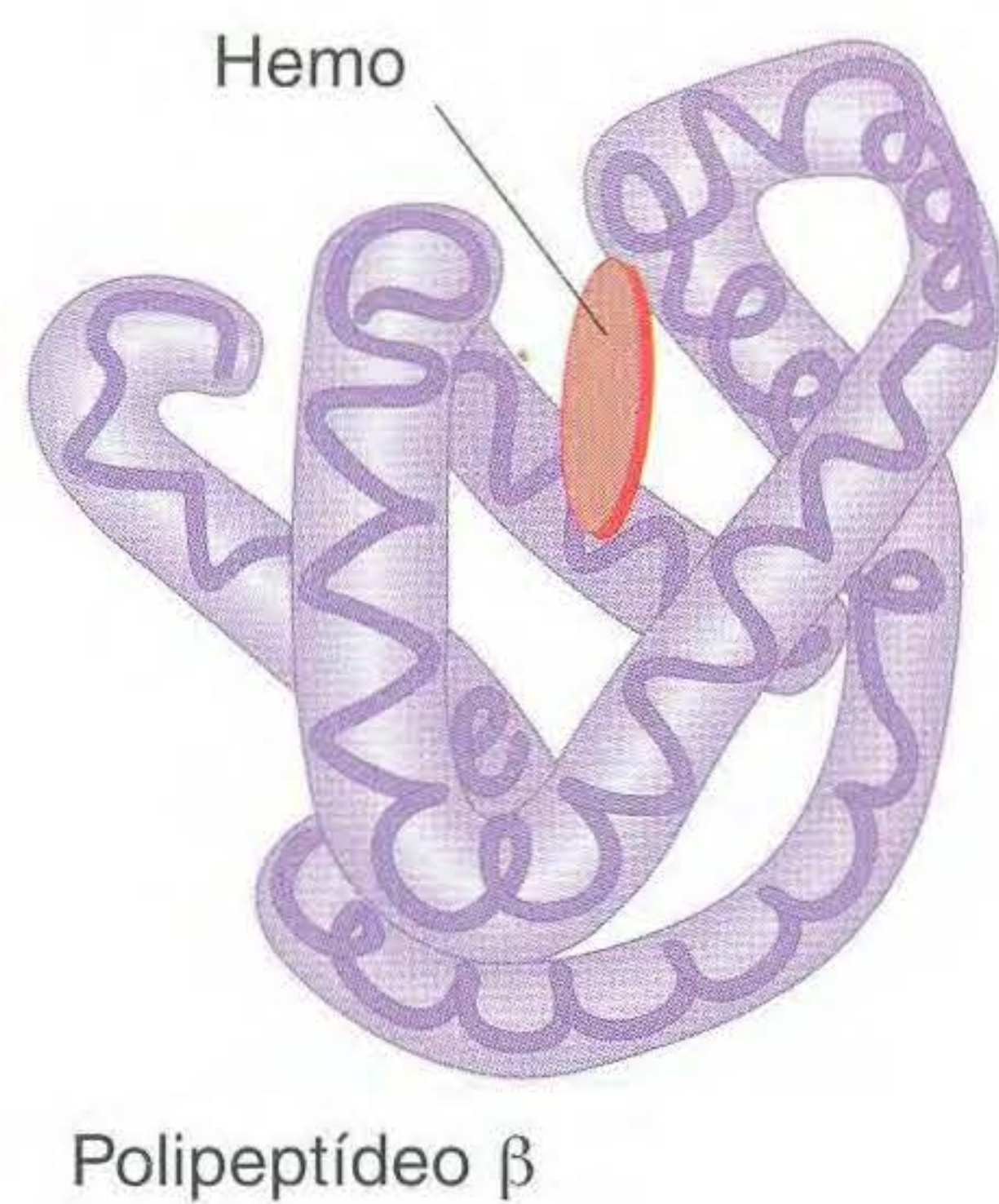
(a) Estrutura primária



(b) Estrutura secundária



(c) Estrutura terciária



(d) Estrutura quaternária

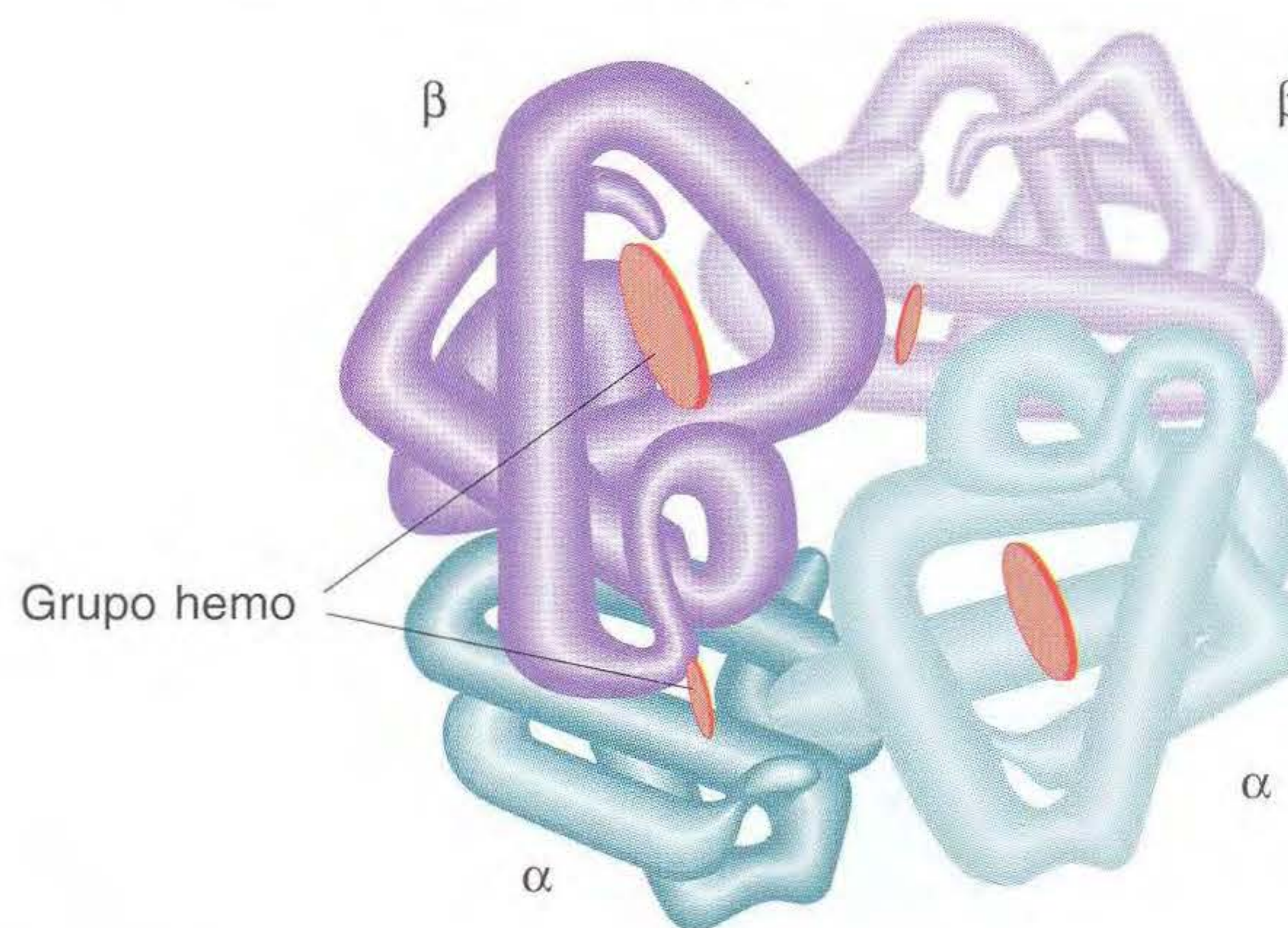


FIG. 9.3 Uma proteína tem quatro níveis de estrutura. (a) Estrutura primária. (b) Estrutura secundária. O polipeptídeo pode formar uma estrutura helicoidal (uma α hélice) ou uma estrutura em zigzag (uma folha pregueada). A folha pregueada tem dois segmentos dispostos em polaridades opostas, como indicado pelas setas. (c) Estrutura terciária. O grupo hemo é uma estrutura em anel não-proteica com um átomo de ferro em seu centro. (d) Estrutura quaternária ilustrada pela hemoglobina, que é composta de quatro subunidades polipeptídicas: duas subunidades α e duas subunidades β .

Muitas proteínas são estruturas compactas; elas são chamadas de **proteínas globulares**. As enzimas e anticorpos estão entre as proteínas globulares mais conhecidas. As proteínas com forma linear, chamadas de **proteínas fibrosas**, são componentes importantes de estruturas tais como a pele, cabelos e tendões.

A forma é importante para uma proteína porque a forma específica de uma proteína permite que ela tenha um papel específico na célula. A forma de uma proteína é determinada por sua sequência primária de aminoácidos e pelas condições na célula que promovem o dobramento e ligação necessários para formar estruturas de maior nível. O dobramento das proteínas em sua conformação correta será discutido no final deste capítulo. A sequência de aminoácidos também determina que grupos R estão presentes em posições específicas e, portanto, disponíveis para se ligar a outros componentes celulares. Os sítios ativos das enzimas são boas ilustrações das interações exatas dos grupos R. Cada enzima tem um bolso chamado de **sítio ativo** ao qual seu substrato ou substratos podem se ajustar. Dentro do sítio ativo, os grupos R de alguns aminoácidos são estrategicamente posicionados para interagir com um substrato e catalisar uma reação química específica.

No momento, as regras pelas quais a estrutura primária é convertida em estrutura de ordem superior são incompletamente compreendidas. Entretanto, pelo conhecimento da sequência primária de uma proteína, podem ser previstas as funções de regiões específicas. Por exemplo, algumas sequências características de proteínas são os pontos de contato com fosfolipídios de membrana que posicionam uma proteína em uma membrana. Outras sequências características atuam para ligar a proteína ao DNA. As sequências de aminoácidos ou proteínas dobradas conservadas que são associadas a determinadas funções são chamadas de **domínios**. Uma proteína pode conter um ou mais domínios separados.

9.2 Colinearidade de Gene e Proteína

A hipótese um gene—uma proteína de Beadle e Tatum (veja o Capítulo 6) foi a fonte da primeira entusiasmada percepção das funções dos genes: os genes, de algum modo, são responsáveis pelo funcionamento das enzimas, e cada gene aparentemente controla uma enzima. Essa hipótese tornou-se um dos grandes conceitos unificantes em biologia, porque forneceu uma ponte que juntou os conceitos e técnicas de pesquisa da Genética e Bioquímica. Quando a estrutura do DNA foi deduzida em 1953, pareceu provável que deve haver uma correspondência linear entre as sequências de nucleotídeos no DNA e as sequências de aminoácidos nas proteínas (tais como uma enzima). Entretanto, só em 1963 é que foi obtida uma demonstração experimental dessa colinearidade. Nesse ano, Charles Yanofsky, da Stanford University, liderou um dos dois grupos de pesquisa que demonstraram a correspondência linear entre os nucleotídeos no DNA e os aminoácidos na proteína.

Yanofsky provou a relação entre os genes alterados e as proteínas alteradas estudando a enzima triptofano sintetase e seu gene. A triptofano sintetase é um heterodímero composto de duas subunidades α e duas β . Ela catalisa a conversão de indol glicerol fosfato em triptofano.



Yanofsky identificou 16 alelos mutantes do gene *trpA* de *E. coli*, cujo alelo tipo selvagem foi conhecido codificando a subunidade α da enzima. Todos os 16 alelos mutantes produzem formas inativas da enzima. Yanofsky foi capaz de mapear a posição do sítio mutante em cada alelo. Em 1963, o sequenciamento do DNA não tinha sido inventado, mas os sítios mutantes podiam ser mapeados pelo método de análise de recombinação, que foi descrito no Capítulo 4.

Pela análise bioquímica da proteína TrpA, Yanofsky mostrou que cada uma das 16 mutações resultou em uma substituição de aminoácidos em uma posição diferente na proteína. Mais entusiasmado, ele mostrou que os sítios mutacionais no mapa gênico do gene *trpA* apareciam na mesma ordem que os correspondentes aminoácidos alterados na cadeia polipeptídica TrpA (Figura 9.4). Além disso, a distância entre os sítios mutantes, medida pela frequência de recombinação, foi demonstrada ser correlata à distância entre os aminoácidos alterados nas proteínas mutantes correspondentes. Assim, Yanofsky demonstrou a **colinearidade**, a correspondência entre a sequência linear do gene e a do polipeptídeo. Subsequentemente, esses resultados foram mostrados aplicando-se a mutações em outras proteínas em geral.

A sequência linear dos nucleotídeos em um gene determina a sequência linear de aminoácidos em uma proteína.

9.3 O Código Genético

Se os genes são segmentos de DNA e se um filamento de DNA é apenas uma fileira de nucleotídeos, então a sequência de nucleotídeos deve, de algum modo, ditar a sequência de aminoácidos nas proteínas. Como a sequência de DNA dita a sequência da proteína? A analogia de um código vem à mente imediatamente. A simples lógica nos diz que, se os nucleotídeos são “letras” em um código, então uma combinação de letras pode formar “palavras” representando aminoácidos diferentes. Primeiro, devemos perguntar como o código é lido. Ele é superposto ou não-superposto? Então, devemos perguntar quantas letras no mRNA constituem uma palavra, ou **códon**, e que códon ou códons representam cada ami-

As seqüências de genes e proteínas são colineares

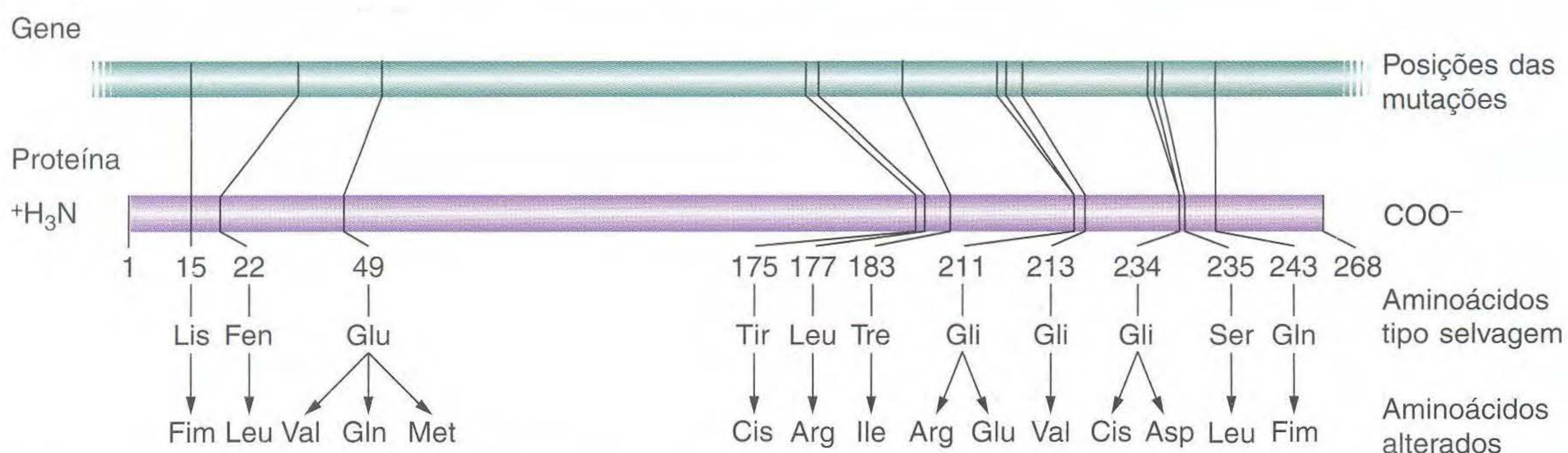


FIG. 9.4 Mutações em *trpA* são colineares com as mudanças de aminoácidos. Embora a ordem das mutações no mapa gênico e a das posições dos aminoácidos sejam as mesmas, as posições relativas diferem porque o mapa gênico é derivado de freqüências de recombinação, que não são uniformes ao longo do tamanho do gene. [De C. Yanofsky, "Gene Structure and Protein Structure." Copyright 1967 by Scientific American. Todos os direitos reservados.]

noácido. O decifrar do código genético é a história contada nesta seção.

Códigos superpostos versus não-superpostos

A Figura 9.5 mostra a diferença entre um código superposto de um não-superposto. O exemplo mostra um código de três letras, ou **trinca**. Para um código não-superposto, aminoácidos consecutivos são especificados por palavras-código consecutivas (códon), como mostrado na parte de baixo da Figura 9.5. Para um código superposto, os aminoácidos consecutivos são especificados por códons que têm algumas bases consecutivas em comum; por exemplo, as últimas duas bases de um códon também podem ser as duas primeiras do códon seguinte. Os códons superpostos são mostrados na parte superior

da Figura 9.5. Assim, para a seqüência AUUGCUCAG em um código não-superposto, as três trincas, AUU, GCU e CAG, codificam os primeiros três aminoácidos, respectivamente. Entretanto, em um código superposto, as trincas AUU, UUG e UGC codificam os primeiros três aminoácidos se a superposição for de duas bases, como mostrado na Figura 9.5.

Em 1961, já estava claro que o código genético era não-superposto. A análise de proteínas mutacionalmente alteradas mostrou que apenas um aminoácido muda de uma vez em uma região da proteína. Esse resultado é previsto por um código não-superposto. Como você pode ver na Figura 9.5, o código superposto prevê que a mudança de uma única base irá alterar até três aminoácidos em posições adjacentes na proteína.

Código superposto versus código genético não-superposto

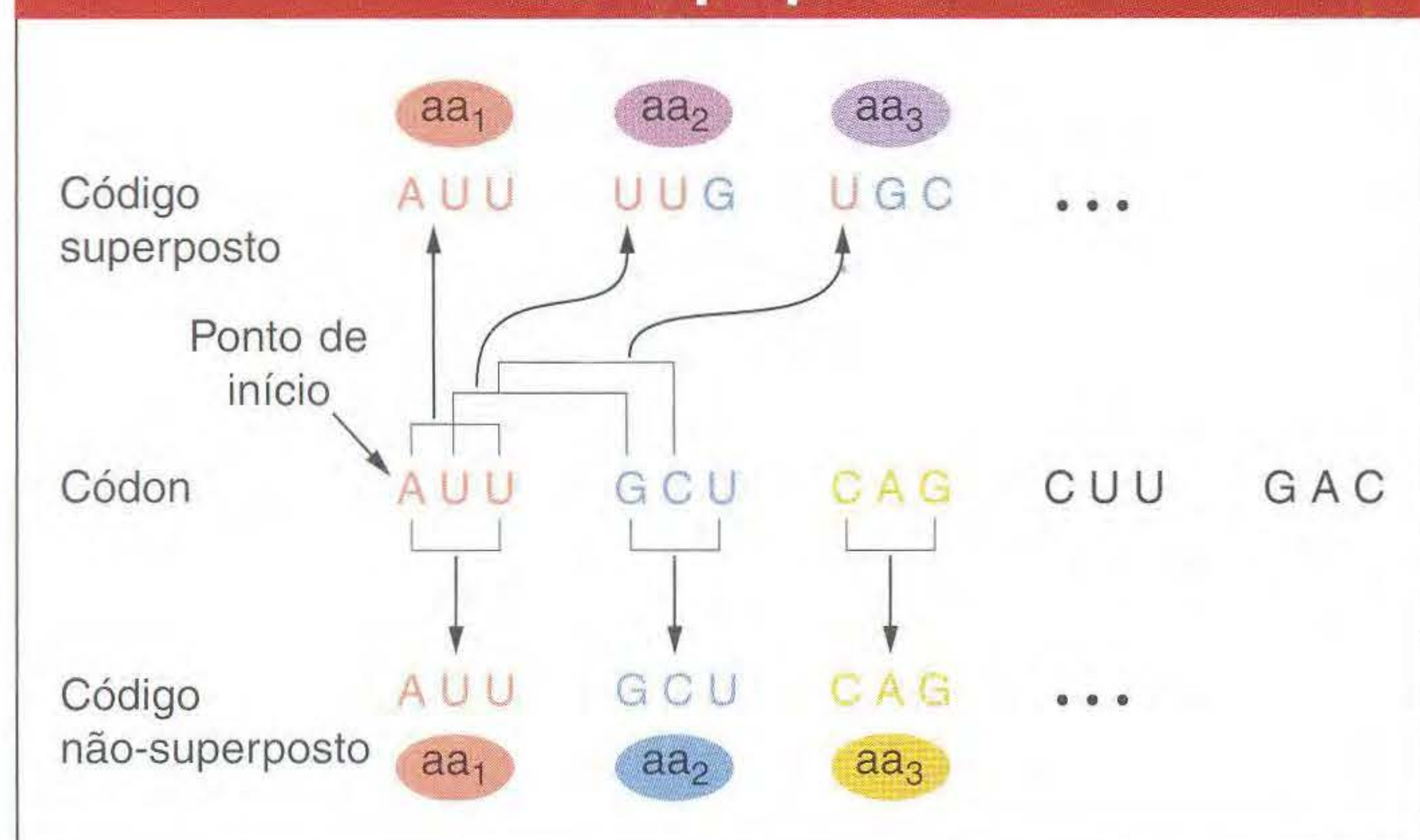


FIG. 9.5 Um código genético superposto e um não-superposto seriam traduzidos de modo diferente em uma seqüência de aminoácidos. O exemplo usa um códon com três nucleotídeos no RNA (um código triplo). Em um código não-superposto, nucleotídeos únicos ocupam posições em vários códons. Nesta ilustração, o terceiro nucleotídeo no RNA, U, é encontrado em três códons. Em um código não-superposto, uma proteína é traduzida lendo seqüencialmente os nucleotídeos em grupos de três. Um nucleotídeo é encontrado apenas em um códon. No exemplo, o terceiro U no RNA é apenas o primeiro códon.

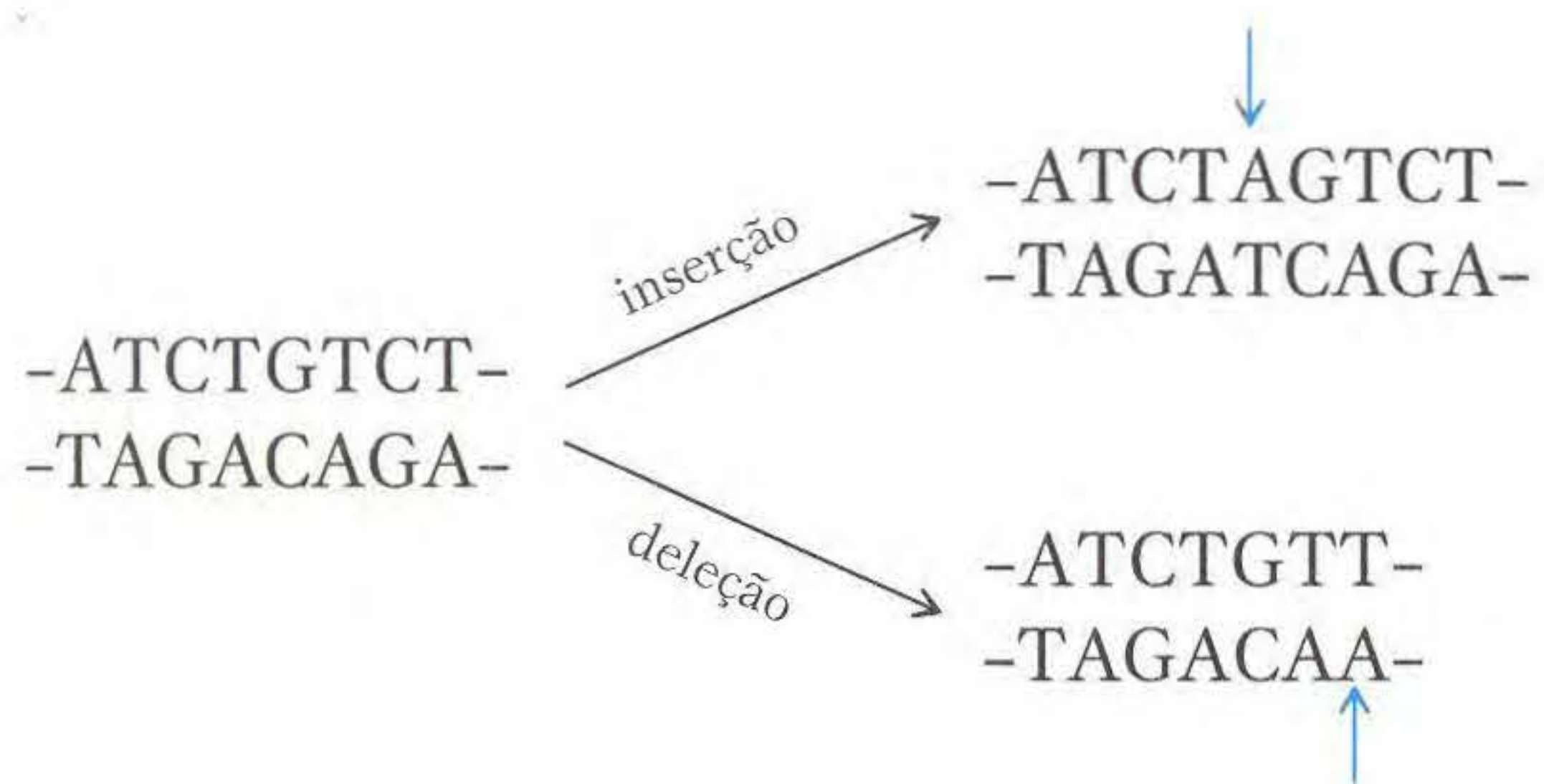
Número de letras no códon

Se uma molécula de mRNA é lida de uma ponta para outra, apenas uma das quatro bases diferentes, A, U, G ou C, pode ser encontrada em cada posição. Assim, se as palavras que codificam aminoácidos tivessem uma letra de tamanho, apenas quatro palavras seriam possíveis. Esse vocabulário não pode ser o código genético, pois precisamos ter uma palavra para cada um dos 20 aminoácidos comumente encontrados nas proteínas celulares. Se as palavras tivessem duas letras de tamanho, então $4 \times 4 = 16$ palavras seriam possíveis; por exemplo, AU, CU ou CC. Esse vocabulário ainda não é grande o suficiente.

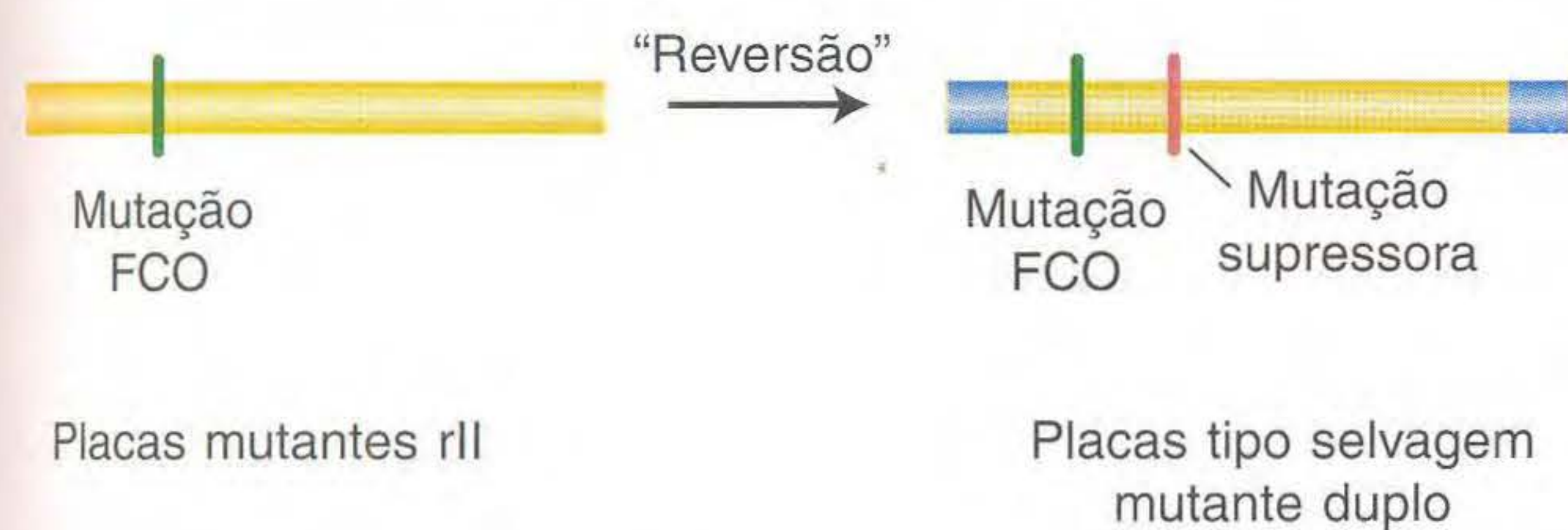
Se as palavras tivessem três letras de tamanho, seriam possíveis $4 \times 4 \times 4 = 64$ palavras; por exemplo, AUU, GCG ou UGC. Esse vocabulário fornece palavras mais do que o suficiente para descrever aminoácidos. Podemos concluir que a palavra código deve consistir em, pelo menos, três nucleotídeos. Entretanto, se todas as palavras são "trincas", então as palavras possíveis são em considerável excesso às 20 necessárias para nomear os aminoácidos comuns. Voltaremos a esse excesso de códons mais tarde no capítulo.

Uso de supressores para demonstrar um código triplo

A prova convincente de que um códon tem, de fato, três letras de tamanho (e não mais que três) veio de belos experimentos genéticos primeiro relatados em 1961 por Francis Crick, Sidney Brenner *et al.* Esses experimentos usaram mutantes no locus *rII* do fago T4. O uso de mutações *rII* na análise de recombinação foi discutido no Capítulo 5. O fago T4 geralmente é capaz de crescer em duas linhagens diferentes de *E. coli*, chamadas B e K. Entretanto, as mutações no gene *rII* mudam a gama de hospedeiros do fago: os fagos mutantes ainda podem crescer em um hospedeiro *E. coli* B, mas não podem crescer em um hospedeiro *E. coli* K. As mutações que causam esse fenótipo *rII* foram induzidas usando uma substância chamada proflavina, que foi tida como agindo pela adição ou deleção de pares únicos de nucleotídeo no DNA. (Essa suposição é baseada em evidências experimentais não apresentadas aqui.) Os exemplos seguintes ilustram a ação da proflavina no DNA bifilamentar.



Começando com uma mutação particular induzida por proflavina chamada FCO, Crick *et al.* encontraram “reversões” (reversões da mutação) que foram capazes de crescer na linhagem K de *E. coli*. A análise genética dessas placas revelou que os “revertentes” não eram idênticos aos tipos selvagens verdadeiros. De fato, a reversão foi encontrada como sendo devida à presença de uma *segunda mutação* em um sítio diferente do de FCO, embora no mesmo gene.



Essa segunda mutação “suprime” a expressão mutante de FCO original. Lembre-se, do Capítulo 6, de que uma mutação supressora contrabalança ou suprime os efeitos de outra mutação.

Como podemos explicar esses resultados? Se for suposto que o gene é lido apenas de uma ponta, então a adição original ou deleção induzida pela proflavina poderia resultar em uma mutação porque ela interrompe um mecanismo normal de leitura que estabelece o grupo de bases a ser lido como palavras. Por exemplo, se cada uma das três bases resultantes no mRNA resultante forma uma palavra, então a “matriz

de leitura” pode ser estabelecida tomando as primeiras três bases do final como a primeira palavra, as três bases seguintes como a segunda palavra, e assim por diante. Nesse caso, uma adição ou deleção induzida por proflavina de um único par no DNA mudaria a matriz de leitura no mRNA a partir desses pontos de correspondência, fazendo com que todas as palavras seguintes fossem lidas erradas. Assim, uma mudança de matriz de leitura pode reduzir a maior parte da mensagem genética a uma confusão sem sentido. Entretanto, a matriz de leitura apropriada pode ser restaurada pela inserção compensatória ou deleção em outra parte, deixando apenas um curto trecho de confusão entre as duas. Considere o seguinte exemplo no qual palavras em inglês com três letras são usadas para representar os códons:

THE FAT CAT ATE THE BIG RAT
Delete C: THE FAT ATA TET HEB IGR AT
Insira A: THE FAT ATA ATE THE BIG RAT

A inserção suprime o efeito da deleção restaurando a maior parte do sentido da frase. Em si, entretanto, a inserção também perturba a frase:

THE FAT CAT AAT ETH EBI GRA T

Se for suposto que o mutante FCO é causado por uma adição, então a segunda mutação (supressora) teria que ser uma deleção, porque, como vimos, apenas uma deleção restauraria a matriz de leitura da mensagem resultante (uma segunda inserção não corrigiria a matriz). Nos diagramas seguintes, usamos uma cadeia hipotética de nucleotídeos para representar simplificada o RNA. Também supomos que as palavras-chave têm três letras de tamanho e são lidas em um sentido (da esquerda para a direita em nossos diagramas).

1. Mensagem tipo selvagem

CAU CAU CAU CAU CAU

2. Mensagem de *rII_a*: As palavras após a adição foram mudadas (×) por mudança de matriz de leitura (palavras marcadas com ✓ não são afetadas).

Adição
CAU ACA UCA UCA UCA U
✓ × × × ×

3. Mensagem *rII_arII_b*: Poucas palavras estão erradas, mas a matriz de leitura é restaurada para as palavras posteriores.

Deleção
CAU ACA UCU CAU CAU
✓ × × ✓ ✓

As poucas palavras erradas no genótipo suprimido poderiam corresponder ao fato de que os “revertentes” (fenótipos suprimidos) que Crick *et al.* recuperaram não se assemelhavam exatamente aos verdadeiros tipos selvagens no fenótipo.

Supusemos aqui que a mudança de matriz original foi uma adição, mas a explicação também serve se for suposto que a mutação original FCO é uma deleção e o supressor é uma adição. Você pode querer verificar por si mesmo. Muito curiosamente, as combinações de *três* adições ou *três* deleções foram mostradas agindo juntas para restaurar um fenótipo tipo selvagem. Essa observação fornece a primeira confirmação experimental de que uma palavra no código genético consiste em três nucleotídeos sucessivos, ou uma trinca. O motivo é que três adições ou três deleções dentro de um gene automaticamente restauram a matriz de leitura no mRNA se as palavras forem trincas.

Redundância do código genético

Como já foi dito, com quatro letras para escolher em cada posição, um códon de três letras poderia fazer $4 \times 4 \times 4 = 64$ palavras. Com apenas 20 palavras necessárias para os 20 aminoácidos comuns, quais são as outras palavras usadas, se alguma? O trabalho de Crick sugeriu que o código genético é **redundante**, significando que cada uma das 64 trincas deveria ter algum significado dentro do código. Para o código ser redundante, alguns dos aminoácidos devem ser especificados por pelo menos duas ou mais trincas diferentes.

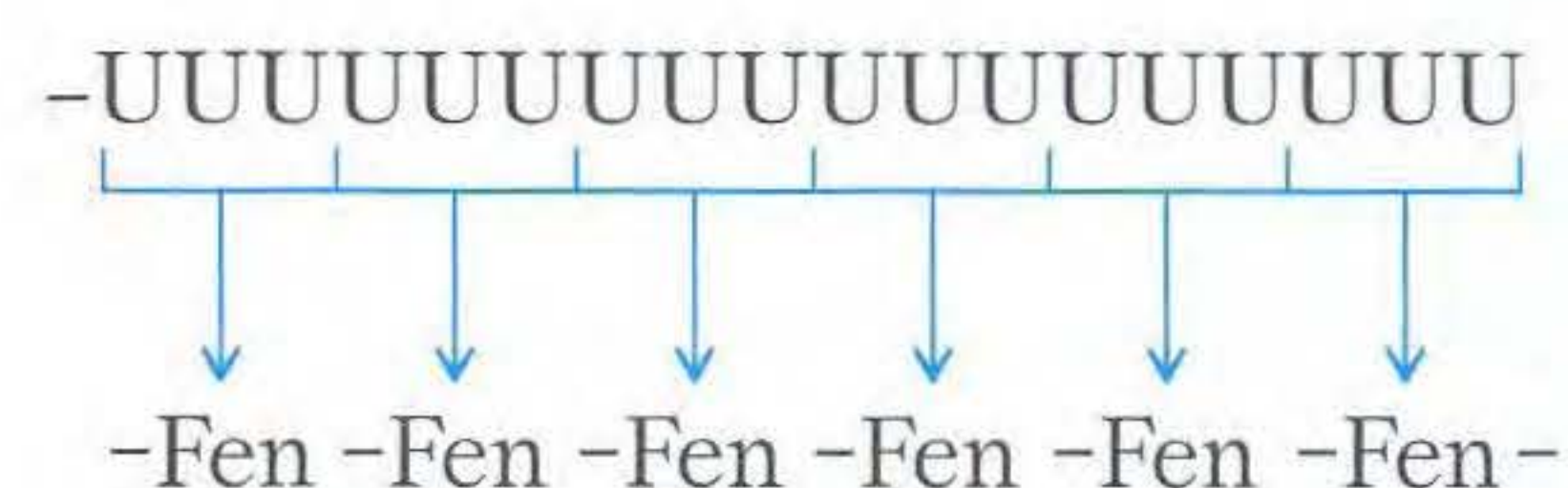
O raciocínio é o seguinte. Se apenas 20 trincas forem usadas, então as outras 44 serão sem sentido, pois não codificariam nenhum aminoácido. Nesse caso, a maioria das mutações de matriz de leitura deveria produzir palavras sem sentido, que supostamente parariam o processo de síntese de proteínas, e a supressão das mudanças de matriz de leitura raramente, se alguma vez, funcionaria. Entretanto, se todas as trincas especificassem algum aminoácido, então as palavras mutadas simplesmente resultariam na inserção de aminoácidos incorretos na proteína. Assim, Crick raciocinou que muitos dos aminoácidos ou todos eles devem ter vários nomes diferentes no código de pares de bases. Essa hipótese depois foi confirmada bioquimicamente.

A discussão até agora demonstrou que

1. O código genético não é superposto.
2. Três bases codificam um aminoácido. Essas trincas são chamadas códons.
3. O código é lido a partir de um ponto inicial fixo e continua até o fim da sequência codificante. Sabemos que o código é lido sequencialmente porque uma única mudança de matriz de leitura em qualquer parte da sequência codificante altera o alinhamento do códon para o resto da sequência.
4. O código é redundante, pois alguns aminoácidos são especificados por mais de um códon.

cas mais fascinantes dos últimos 50 anos. Após as técnicas experimentais tornarem-se disponíveis, o código genético foi rapidamente decifrado.

O ponto importante foi a descoberta de como fazer um mRNA sintético. Se os nucleotídeos do RNA são misturados com uma enzima especial (polinucleotídeo fosforilase), é formado um único filamento de RNA na reação. Ao contrário da transcrição, nenhum molde de DNA é necessário para essa síntese, e, assim, os nucleotídeos são incorporados aleatoriamente. A habilidade em sintetizar RNA ofereceu uma perspectiva excitante de criar sequências específicas de mRNA e, então, ver quais aminoácidos elas especificariam. O primeiro mensageiro sintético obtido foi feito misturando apenas nucleotídeos uracil com a enzima de síntese de RNA, produzindo ...UUUU... [poli(U)]. Em 1961, Marshall Nirenberg e Heinrich Matthaei misturaram poli(U) com a maquinaria de síntese de proteínas de *E. coli in vitro* e observaram a formação de uma proteína. O principal entusiasmo estava centrado na questão da sequência de aminoácidos dessa proteína. Ficou demonstrado ser uma polifenilalanina — uma fileira de moléculas de fenilalanina ligadas para formar um polipeptídeo. Assim, a trinca UUU deve codificar a fenilalanina:



Por essa descoberta, Nirenberg recebeu o Prêmio Nobel.

Em seguida, os mRNA contendo dois tipos de nucleotídeos em grupos repetidos foram sintetizados. Por exemplo, o mRNA sintético tendo a sequência $(AGA)_n$, que é uma sequência longa de AGAAGAAGAAGAAGA, foi usada para estimular a síntese de polipeptídeos *in vitro* (em um tubo de ensaio que também continha um extrato celular com todos os componentes necessários para a tradução). As sequências de polipeptídeos resultantes foi observada de uma variedade de tais testes, com o uso de trincas diferentes que residem em outros RNA sintéticos. A partir desses testes, muitas palavras-códon puderam ser verificadas. (Esse tipo de experimento é detalhado no Problema 33 ao final deste capítulo. Para resolvê-lo, você pode colocar-se no lugar de H. Gobind Khorana, que recebeu um Prêmio Nobel por dirigir os experimentos.)

Enfoques experimentais adicionais levaram à ligação de cada aminoácido a um ou mais códons. Lembre que o código foi proposto como sendo redundante, significando que alguns aminoácidos tinham mais de um códon determinante. Essa redundância pode ser claramente vista na Figura 9.6, que dá os códons e os aminoácidos que eles especificam. Praticamente todos os organismos em nosso planeta usam o mesmo código genético. (Existem apenas algumas exceções nas quais um pequeno número de códons têm significados diferentes — por exemplo, nos genomas mitocondriais.)

Códons de fim

Você deve ter notado, na Figura 9.6, que alguns códons não especificam nenhum aminoácido. Esses códons são códons

Decifrando o código

Decifrar o código genético, a determinação do aminoácido especificado por cada trinca, foi uma das conquistas genéti-

O código genético

		Segunda letra				
		U	C	A	G	
Primeira letra	U	UUU } Fen UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tir UAC } UAA } Fim UAG } Fim	UGU } Cis UGC } UGA } Fim UGG } Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG } Met	ACU } ACC } Tre ACA } ACG }	AAU } Asn AAC } AAA } Lis AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gli GGA } GGG }	U C A G

FIG. 9.6 O código genético designa os aminoácidos especificados por cada códon.

de fim, ou de término. Eles podem ser vistos como sendo similares a períodos ou vírgulas pontuando a mensagem codificada no DNA.

Uma das primeiras indicações da existência de códons de fim veio em 1965, do trabalho de Brenner com o fago T4. Brenner analisou algumas mutações (m_1 - m_6) em um único gene que controla a proteína da cabeça do fago. Ele observou que a proteína da cabeça de cada mutante era uma cadeia polipeptídica mais curta do que a do tipo selvagem. Brenner examinou as pontas das proteínas encurtadas e comparou-as à proteína tipo selvagem. Para cada mutante, ele registrou o aminoácido seguinte que *teria* sido inserido para continuar a cadeia tipo selvagem. Os aminoácidos para as seis mutações eram glutamina, lisina, ácido glutâmico, tirosina, triptofano e serina. Esses resultados não apresentam imediatamente um padrão óbvio, mas Brenner brilhantemente deduziu que alguns códons para esses aminoácidos eram similares. Especificamente, cada um desses códons pode mutar para o códon UAG por uma única mudança em um par de nucleotídeos do DNA. Ele, portanto, postulou que UAG é um códon de fim (término) — um sinal para o mecanismo de tradução de que a proteína está agora completa.

UAG foi o primeiro códon decifrado; ele é chamado de códon âmbar. Os mutantes que são defeituosos devido à presença de um códon âmbar anormal são chamados de *mutantes âmbar*. Os outros códons de fim são UGA e UAA. Analogamente ao códon âmbar, UGA é chamado de códon opala e UAA é chamado de códon ocre. Os mutantes que são defeituosos porque contêm códons opala e ocre anormais são chamados de *mutantes opala e ocre*, respectivamente. Os códons de fim são geralmente chamados de códons sem sentido porque não designam aminoácidos.

Além de uma proteína mais curta de cabeça, os fagos mutantes de Brenner tinham outra característica interessante

em comum: a presença de uma mutação supressora (*su*-) no cromossomo hospedeiro faria com que o fago desenvolvesse uma proteína de cabeça de tamanho normal de cadeia (tipo selvagem) a despeito da presença da mutação *m*. Devemos considerar os códons de fim e seus supressores mais adiante, após lidarmos com o processo de síntese de proteínas.

9.4 tRNA: o Adaptador

Após tornar-se conhecido que a sequência de aminoácidos de uma proteína foi determinada pelos códons de trinca no mRNA, os cientistas começaram a imaginar como era feita essa determinação. Um modelo inicial, rapidamente descartado como muito simples e improvável, propôs que os códons do mRNA poderiam dobrar-se e formar 20 cavidades distintas que se ligam diretamente a aminoácidos na ordem correta. Em vez disso, em 1958, Crick reconheceu que

É, portanto, uma hipótese natural que o aminoácido seja levado para o molde por uma molécula adaptadora, e que o adaptador é a parte que se ajusta ao RNA. Em sua forma mais simples [esta hipótese] seriam necessários 20 adaptadores, um para cada aminoácido.

Ele especulou que o adaptador “pode conter nucleotídeos. Isso lhes permitiria juntar-se ao molde de RNA pelo mesmo ‘pareamento’ de bases encontrado no DNA”. Além disso, “uma enzima separada seria necessária para juntar cada adaptador a seu próprio aminoácido”.

Hoje sabemos que a “hipótese do adaptador” de Crick está amplamente correta. Os aminoácidos de fato estão ligados a um adaptador (lembre que os adaptadores constituem uma classe especial de RNA estáveis: os chamados *RNA transportadores*). Cada aminoácido torna-se ligado a um tRNA específico, que então leva esse aminoácido para o ribossomo, o complexo molecular que irá ligar o aminoácido a um polipeptídeo crescente.

Tradução do códon pelo tRNA

A estrutura do tRNA possui o segredo da especificidade entre um códon do mRNA e o aminoácido que ele designa. A molécula unifilamentar de tRNA tem uma forma de trevo consistindo em quatro hastes de dupla hélice e três alças unifilamentares (Figura 9.7a). A alça do meio de cada tRNA é chamada de alça do anticódon porque leva uma trinca de nucleotídeos chamada de **anticódon**. Essa sequência é complementar ao códon para o aminoácido levado pelo tRNA. O anticódon no tRNA e o códon no mRNA ligam-se a um pareamento específico de bases RNA com RNA. (Novamente, vemos funcionando o princípio da complementariedade de ácidos nucleicos, dessa vez na ligação de dois RNA diferentes.) Como os códons no mRNA são lidos no sentido 5' → 3', os anticódons são orientados e escritos no sentido 3' → 5', como mostra a Figura 9.7a.

Os aminoácidos são ligados aos tRNA por enzimas chamadas de **aminoacil-tRNA sintetases**. Existem 20 dessas marcantes enzimas na célula, uma para cada um dos 20 aminoácidos. O tRNA com um aminoácido ligado é dito **carre-**

A estrutura do RNA transportador

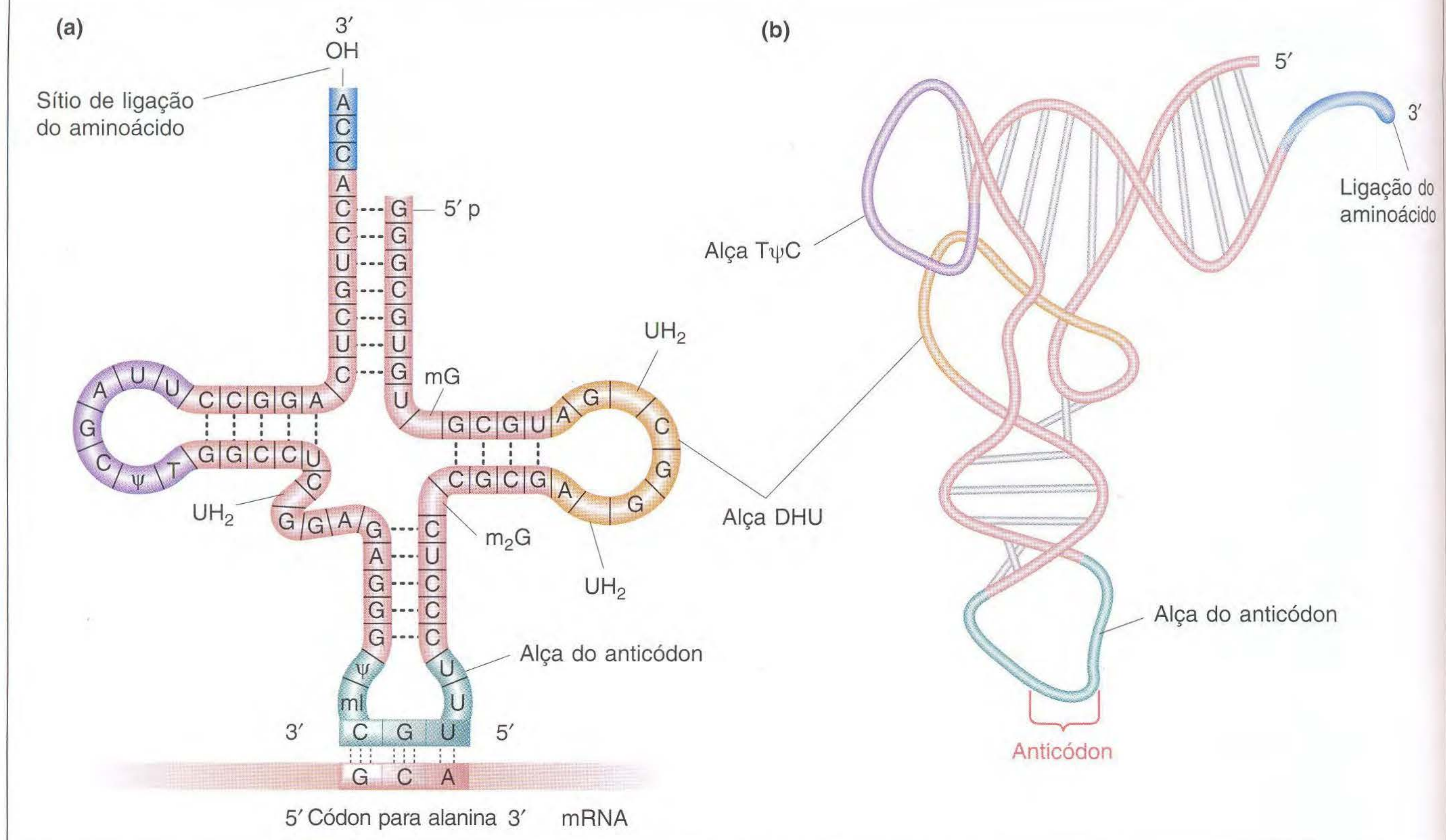


FIG. 9.7 (a) A estrutura do tRNA de alanina de levedura, mostrando o anticódon do tRNA ligando-se ao códon complementar no mRNA. (b) Diagrama da estrutura tridimensional do tRNA de fenilalanina de levedura. As abreviações ψ , mG, m₂G, ml e UH₂ referem-se às bases modificadas pseudo-uridina, metilguanósina, dimetilguanósina, metilinosina e diidrouridina, respectivamente. [(a) De S. Arnott, "The Structure of Transfer RNA," *Prog. Biophys. Mol. Biol.* 22, 1971, 186; (b) de L. Stryer, *Biochemistry*, 4th ed. Copyright 1995 by Lubert Stryer. A parte b é baseada em um desenho de Sung-Hou Kim.]

Uma aminoacil-tRNA sintetase liga um aminoácido a seu tRNA

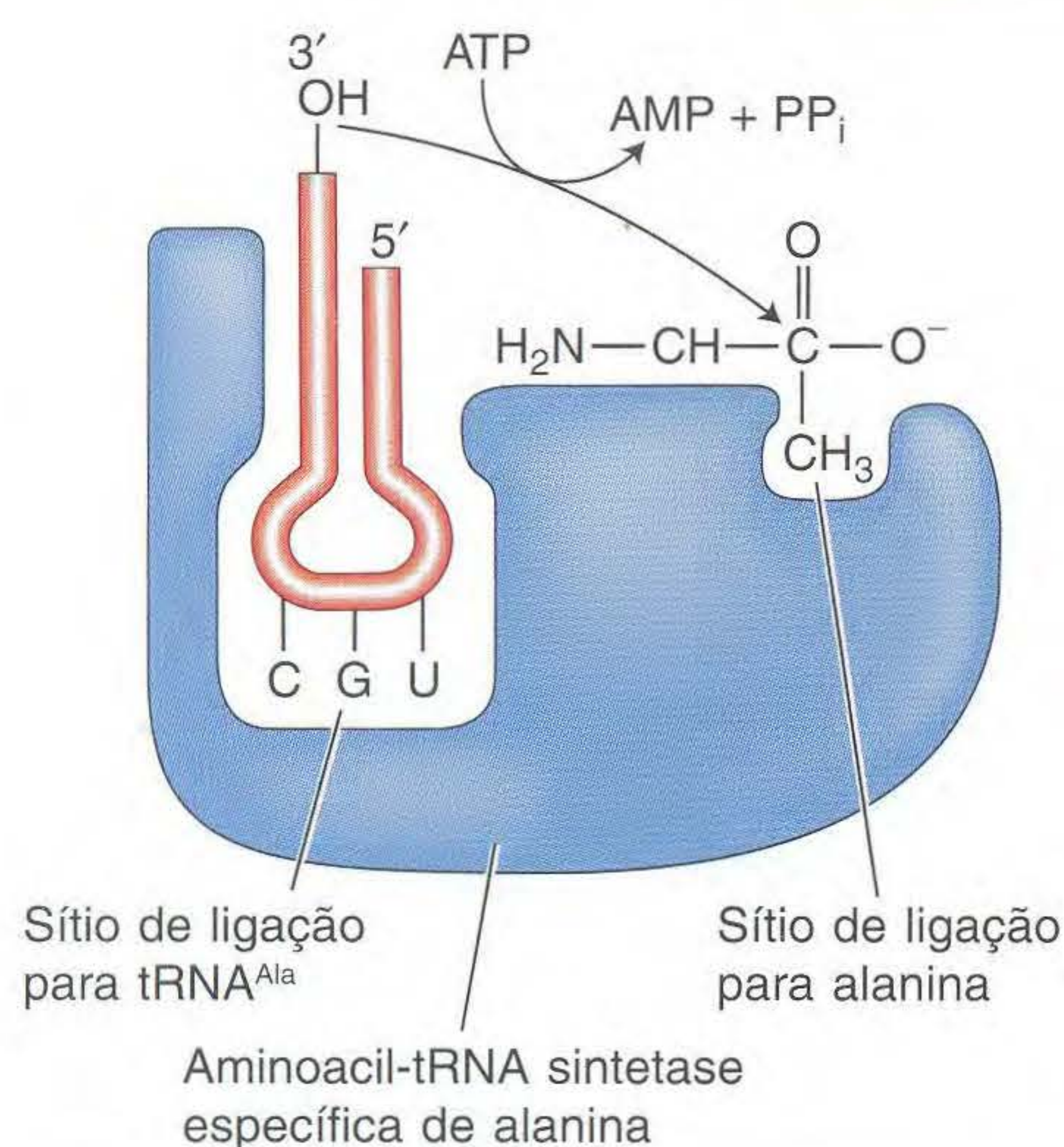


FIG. 9.8 Cada aminoacil-tRNA sintetase tem bolsos de ligação para um aminoácido específico e seu tRNA cognato. Desse modo, um aminoácido é covalentemente ligado ao tRNA com o anticódon correspondente.

gado. Cada aminoácido tem uma sintetase específica que o liga apenas aos tRNA que reconhecem os códons para esse determinado aminoácido. Para catalisar essa reação, as sintetases têm dois sítios de ligação: um para o aminoácido e o outro para seu tRNA cognato (Figura 9.8). Um aminoácido é ligado à ponta 3' de seu tRNA, o aminoácido alanina no caso mostrado nas Figuras 9.7 e 9.8.

O trevo "achatado" mostrado na Figura 9.7a não é a conformação normal das moléculas de tRNA; um tRNA normalmente existe como um trevo dobrado em L, como mostrado na Figura 9.7b. A estrutura tridimensional do tRNA foi determinada com o uso de cristalografia de raios X. Nos anos em que foi usada para deduzir a estrutura da dupla hélice do DNA, essa técnica foi refinada, de modo que agora pode ser usada para determinar a estrutura de macromoléculas complexas tais como o ribossomo. Embora os tRNA difiram em sua sequência primária de nucleotídeos, todos os RNA dobras-se praticamente na mesma conformação em L, exceto por diferenças na alça de anticódon e ponta aminoacil. Essa similaridade de estrutura pode facilmente ser vista na Figura 9.9, que mostra dois tRNA diferentes superpostos. A conservação da estrutura nos diz que a forma é importante para o funcionamento do tRNA.

Dois tRNA superpostos

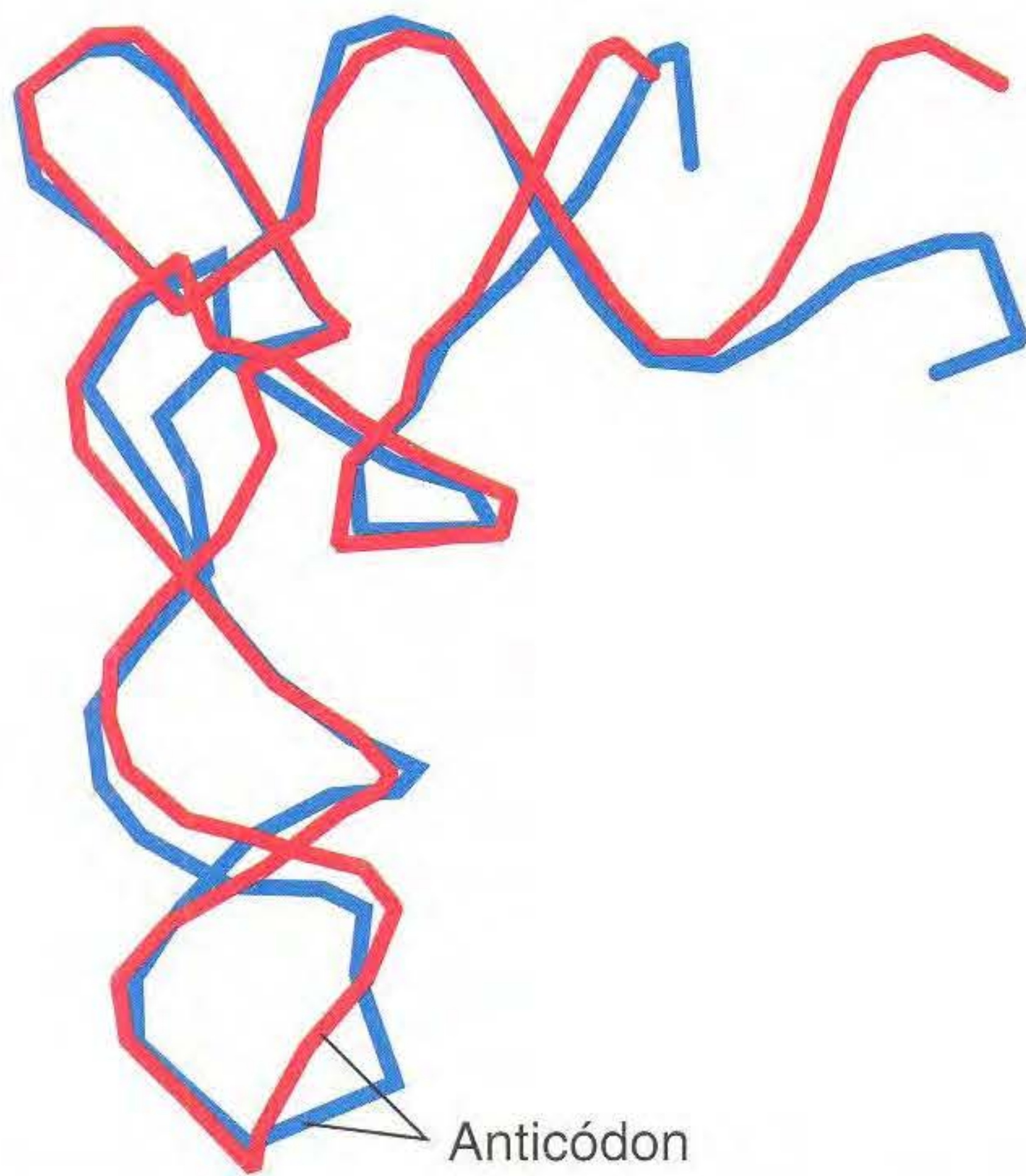
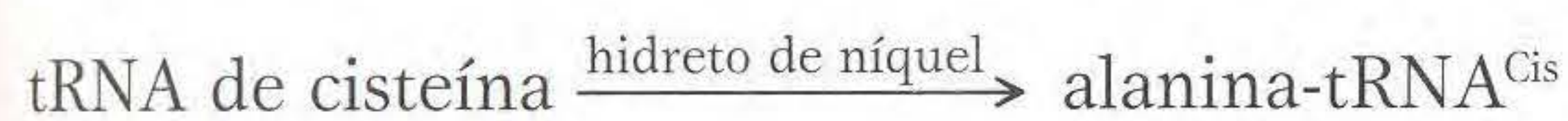


FIG. 9.9 Quando dobrado em suas estruturas tridimensionais corretas, o tRNA de levedura para glutamina (azul) superpõe quase totalmente o tRNA de levedura para fenilalanina (vermelho), exceto para a alça de anticódon e ponta aminoacil. [De M.A. Rould, J.J. Perona, D. Soll, and T.A. Steitz, "Structure of *E. coli* Glutamyl-tRNA Synthetase Complexed with tRNA(Gln) and ATP in 2,8 Å of Resolution," *Science* 246, 1989, 1135–1142.]

O que ocorreria se o aminoácido errado fosse covalentemente ligado a um tRNA? Um experimento convincente responde a essa pergunta. O experimento usou cisteinil-tRNA ($tRNA^{Cis}$), o tRNA específico para cisteína. Esse tRNA foi "carregado" com cisteína, significando que cisteína foi ligada ao tRNA. O tRNA carregado foi tratado com hidreto de níquel, que converte a cisteína (enquanto ainda ligada ao $tRNA^{Cis}$) em outro aminoácido, alanina, sem afetar o tRNA:



A síntese de proteínas com essa espécie híbrida tem alanina sempre onde esperávamos cisteína. O experimento demonstrou que os aminoácidos são "analfabetos"; eles são inseridos na posição apropriada porque os tRNA "adaptadores" reconhecem os códons do mRNA e inserem apropriadamente seus aminoácidos ligados. Assim, a ligação do aminoácido com seu tRNA congênito é uma etapa crítica para garantir que a proteína seja corretamente sintetizada. Se o aminoácido errado for ligado, não há meio de evitar que ele seja incorporado a uma cadeia proteica crescente.

Novamente a redundância

Como pode ser visto na Figura 9.6, o número de códons para um único aminoácido varia, indo de um códon (UGG para triptofano) até seis (UCC, UCU, UCA, UCG, AGC ou AGU para serina). Por que o código genético contém essa variação ainda não está claro, mas dois fatores respondem por ela:

1. A maioria dos animais pode ser levada aos ribossomos por vários tipos de tRNA alternativos. Cada tipo tem um anticódon diferente que faz par de bases com um códon diferente no mRNA.

A oscilação permite que um tRNA reconheça dois códons

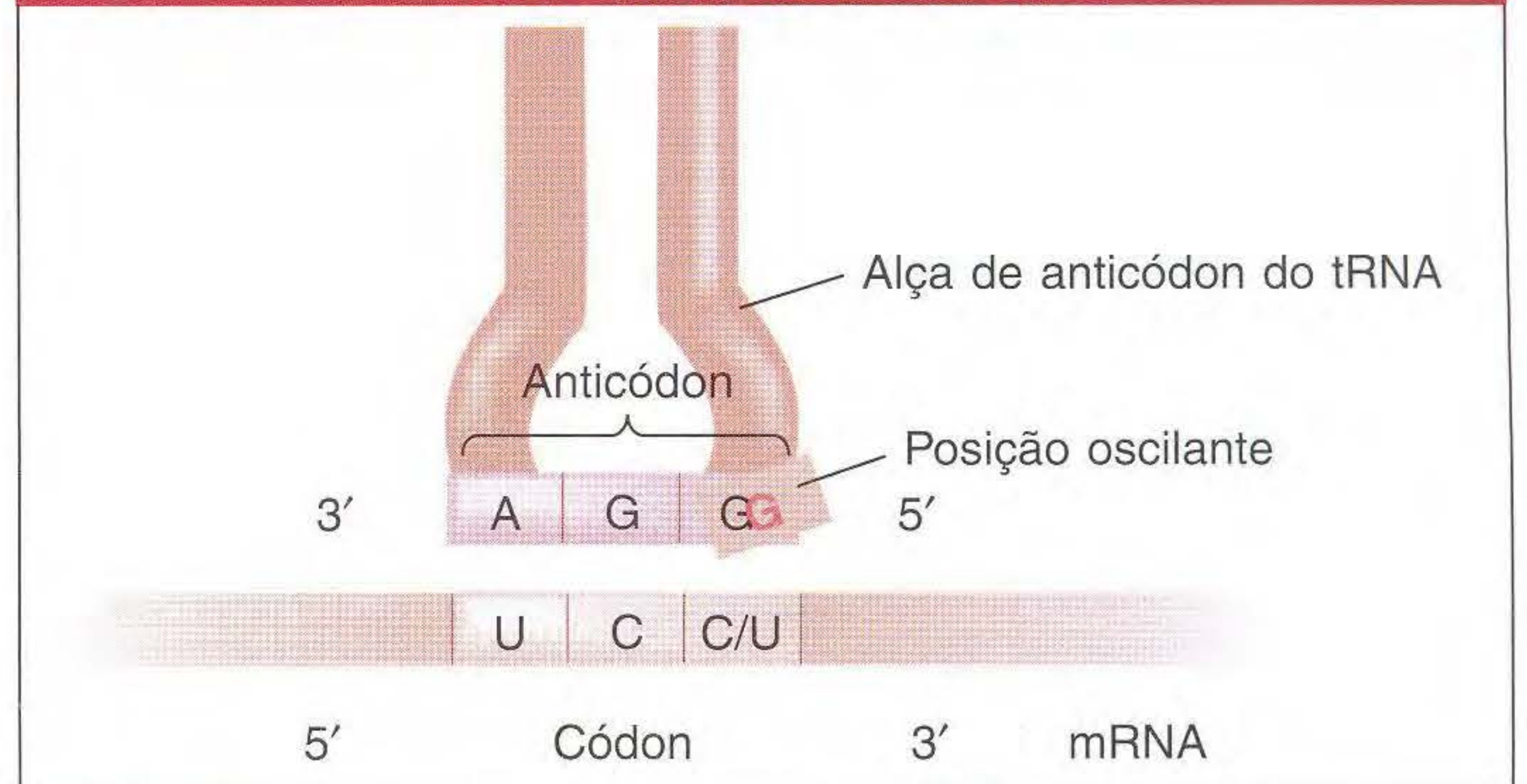


FIG. 9.10 No terceiro sítio (ponta 5') do anticódon, G pode ter uma das duas posições de oscilação, sendo assim capaz de parear com U ou C. Essa habilidade significa que uma única espécie de tRNA levando um aminoácido (nesse caso, serina) pode reconhecer dois códons – UCU e UCC – no mRNA.

2. Algumas espécies de tRNA carregados podem trazer seus aminoácidos específicos para qualquer um de vários códons. Esses tRNA reconhecem e ligam-se a vários códons alternativos, não apenas a um com uma sequência complementar, por um tipo de pareamento de bases fraco na ponta 3' do códon e ponta 5' do anticódon. Esse pareamento fraco é chamado de **oscilação (wobble)**.

A oscilação é uma situação na qual o terceiro nucleotídeo de um anticódon (na ponta 5') pode formar dois alinhamentos (Figura 9.10). Esse terceiro pode formar pontes de hidrogênio seja com seu nucleotídeo complementar normal na terceira posição do códon, seja com um nucleotídeo diferente nessa posição. As "regras da oscilação" ditam que nucleotídeos podem e não podem formar pontes de hidrogênio com nucleotídeos alternativos pela oscilação (Quadro 9.1). No Quadro 9.1, a letra *I* representa inosina, uma das raras bases encontradas no tRNA, geralmente no anticódon.

O Quadro 9.2 cita todos os códons para serina e mostra como três tRNA diferentes ($tRNA^{Ser_1}$, $tRNA^{Ser_2}$ e $tRNA^{Ser_3}$) podem parear com esses códons. Alguns organismos possuem uma espécie adicional de tRNA (que podemos representar como $tRNA^{Ser_4}$) que tem um anticódon idêntico ao dos três anticódons mostrados no Quadro 9.2, mas difere em

Quadro 9.1 Pareamentos Códon-Anticódon Permitidos pelas Regras da Oscilação

Ponta 5' do anticódon	Ponta 3' do códon
G	C ou U
C	só G
A	só U
U	A ou G
I	U, C ou A

Quadro 9.2 Diferentes tRNA que Podem Servir de Códon de Serina

tRNA	Anticódon	Códon
tRNA ^{Ser} ₁	ACG + oscilação	UCC
		UCU
tRNA ^{Ser} ₂	AGU + oscilação	UCA
		UCG
tRNA ^{Ser} ₃	UCG + oscilação	AGC
		AGU

sua seqüência de nucleotídeos em outras partes da molécula. Esses quatro tRNA são chamados de **tRNA isoaceptores** porque aceitam os mesmos aminoácidos, mas são transcritos de diferentes genes de tRNA.

O código genético é dito como redundante porque, em muitos casos, mais de um códon é atribuído a um único aminoácido; além disso, vários códon podem parear com mais de um anticódon (oscilação).

► 9.5 Ribossomos

A síntese de proteínas ocorre quando o tRNA e as moléculas de mRNA se associam aos *ribossomos*. A tarefa dos tRNA e ribossomo é traduzir a seqüência de códon de nucleotídeos em mRNA em uma seqüência de aminoácidos na proteína. O termo *máquina biológica* foi usado nos capítulos precedentes para caracterizar complexos multissubunitários que desempenham funções celulares. O replissomo, por exemplo, é uma máquina biológica que pode replicar o DNA com precisão e velocidade. O sítio de síntese de proteínas, o ribossomo, é muito maior e mais complexo do que as máquinas descritas até agora. Sua complexidade é devida ao fato de que ele desempenha várias tarefas com precisão e velocidade. Por esse motivo, é melhor pensar no ribossomo como uma fábrica contendo muitas máquinas que atuam em conjunto. Vejamos como essa máquina é organizada para desempenhar suas várias funções.

Em todos os organismos, o ribossomo consiste em uma subunidade pequena e uma grande, cada uma feita de RNA (chamado de RNA ribossômico ou rRNA) e proteínas. Cada subunidade é composta de um a três tipos de rRNA e até 50 proteínas. As subunidades ribossômicas foram originalmente caracterizadas por sua velocidade de sedimentação quando giradas em uma ultracentrífuga, e, assim, seus nomes são derivados de seus coeficientes de sedimentação em unidades Svedberg (S), que é uma indicação do tamanho molecular. Em procariontes, as subunidades pequena e grande são chamadas de 30S e 50S, respectivamente, e se associam para formar uma partícula 70S (Figura 9.11a). As contrapartes eucarióticas são chamadas de 40S e 60S, com 80S para o ribossomo completo (Figura 9.11b). Embora os ribossomos eucarióticos sejam maiores devido a seus componentes numerosos e maiores, os componentes e as etapas na síntese de proteínas são, em geral, similares. As similaridades indicam claramente

que a tradução é um processo antigo que se originou em um ancestral comum dos eucariontes e procariontes.

Quando os ribossomos foram primeiro estudados, o fato de que quase dois terços de sua massa seja de RNA e apenas um terço de proteínas foi surpreendente. Por décadas, os rRNA foram supostos funcionando como o arcabouço necessário para a montagem correta das proteínas ribossomais. Esse papel parecia lógico porque os RNA dobram-se por pareamento intramolecular de bases em estruturas secundárias estáveis (Figura 9.12). De acordo com esse modelo, as proteínas ribossômicas eram apenas responsáveis por efetuar as etapas importantes na síntese de proteínas. Essa visão mudou com a descoberta, nos anos de 1980, de RNA catalíticos (veja o Capítulo 8). Como você verá, os cientistas agora acreditam que os rRNA, ajudados por proteínas ribossômicas, efetuam a maioria das etapas importantes da síntese de proteínas.

Características dos ribossomos

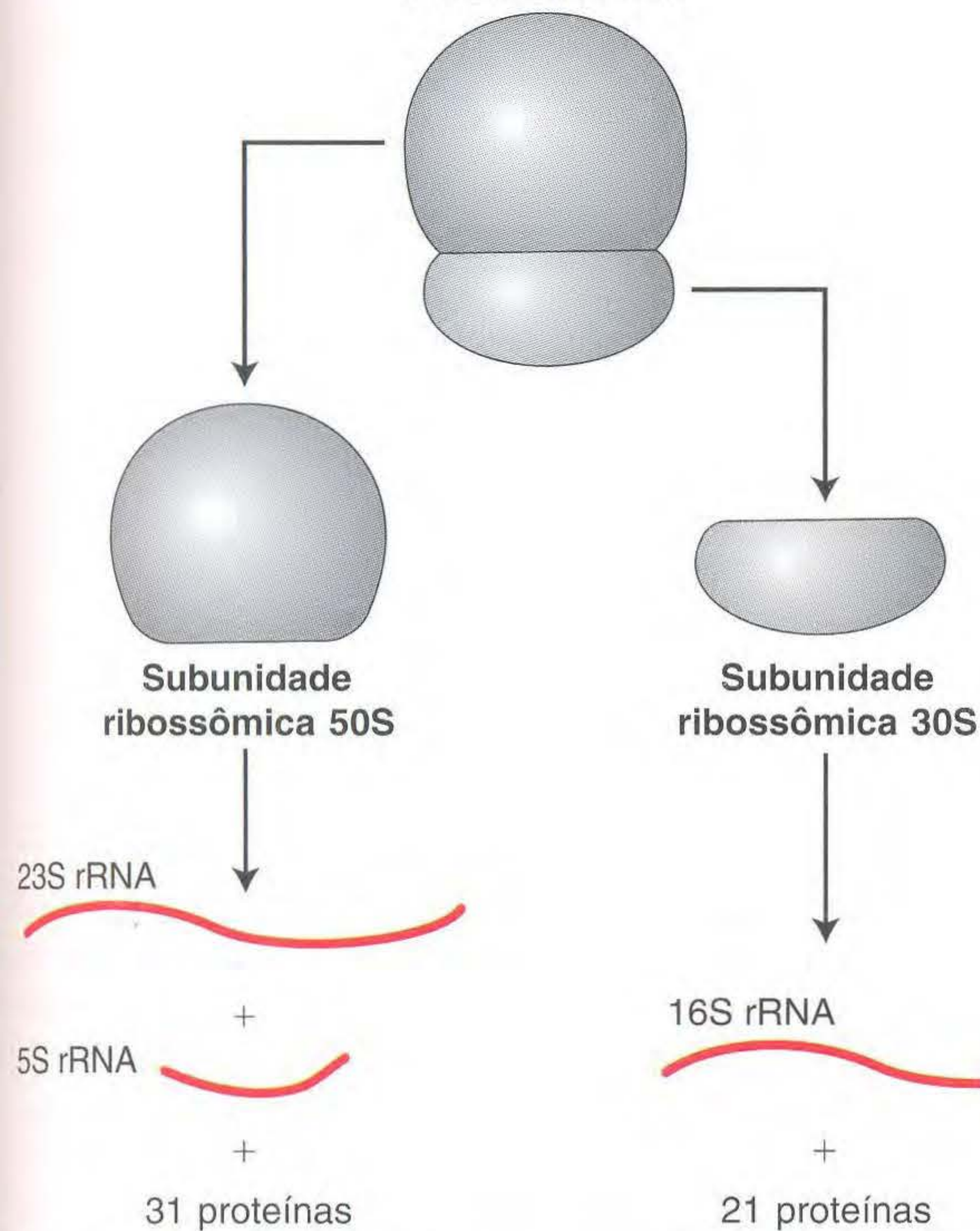
Os ribossomos unem os outros dois participantes importantes na síntese de proteínas — as moléculas de tRNA e mRNA — para traduzir a seqüência de nucleotídeos de um mRNA em uma seqüência de aminoácidos de uma proteína. As moléculas de tRNA e mRNA estão posicionadas no ribossomo de modo que o códon do mRNA possa interagir com o anticódon do tRNA. Os sítios principais de interação são ilustrados na Figura 9.13. O sítio de ligação ao mRNA está totalmente dentro da subunidade menor. Existem três sítios de ligação para as moléculas de tRNA. Cada tRNA ligado une as subunidades 30S e 50S, com sua ponta de anticódon na primeira e sua ponta aminoacil (levando o aminoácido) na última. O **sítio A** (para aminoacil) liga um aminoacil-tRNA que chega cujo anticódon corresponde ao códon no sítio A da subunidade 30S. À medida que continuamos no sentido 5' do mRNA, o códon seguinte interage com o anticódon do tRNA no **sítio P** (para peptidil) da subunidade 30S. O tRNA no sítio P liga-se à cadeia polipeptídica crescente, parte da qual entra em uma estrutura tipo túnel na subunidade 50S. O **sítio E** (de saída, *exit*) contém um tRNA desacilado (ele não leva mais um aminoácido) que está pronto para ser liberado do ribossomo. Ainda não está claro se as interações códon-anticódon também tomam parte entre o mRNA e o tRNA no sítio E.

Duas regiões adicionais no ribossomo são críticas para a síntese de proteínas. O **centro decodificador**, na subunidade de 30S, garante que apenas os tRNA levando anticódons que se ajustam aos códon (chamados de tRNA *cognatos*) serão aceitos no sítio A. Os tRNA cognatos associam-se ao **centro peptidiltransferase** na subunidade 50S, onde a formação da ligação peptídica é catalisada. Recentemente, muitos laboratórios, especialmente o de Thomas Steitz, usaram cristalografia de raios X para “resolver” a estrutura do ribossomo na associação com tRNA ao nível atômico. Os resultados desses elegantes estudos mostraram claramente que ambos os centros são compostos inteiramente de regiões do rRNA; isto é, os contatos importantes nesses centros são contatos tRNA-rRNA. A formação de ligações peptídicas é ainda tida como catalisada pelo sítio ativo no RNA ribossômico, sendo

Proteína e moléculas de RNA compõem as duas subunidades de um ribossomo

(a) Procariótico

Ribossomo 70S



(b) Eucariótico

Ribossomo 80S

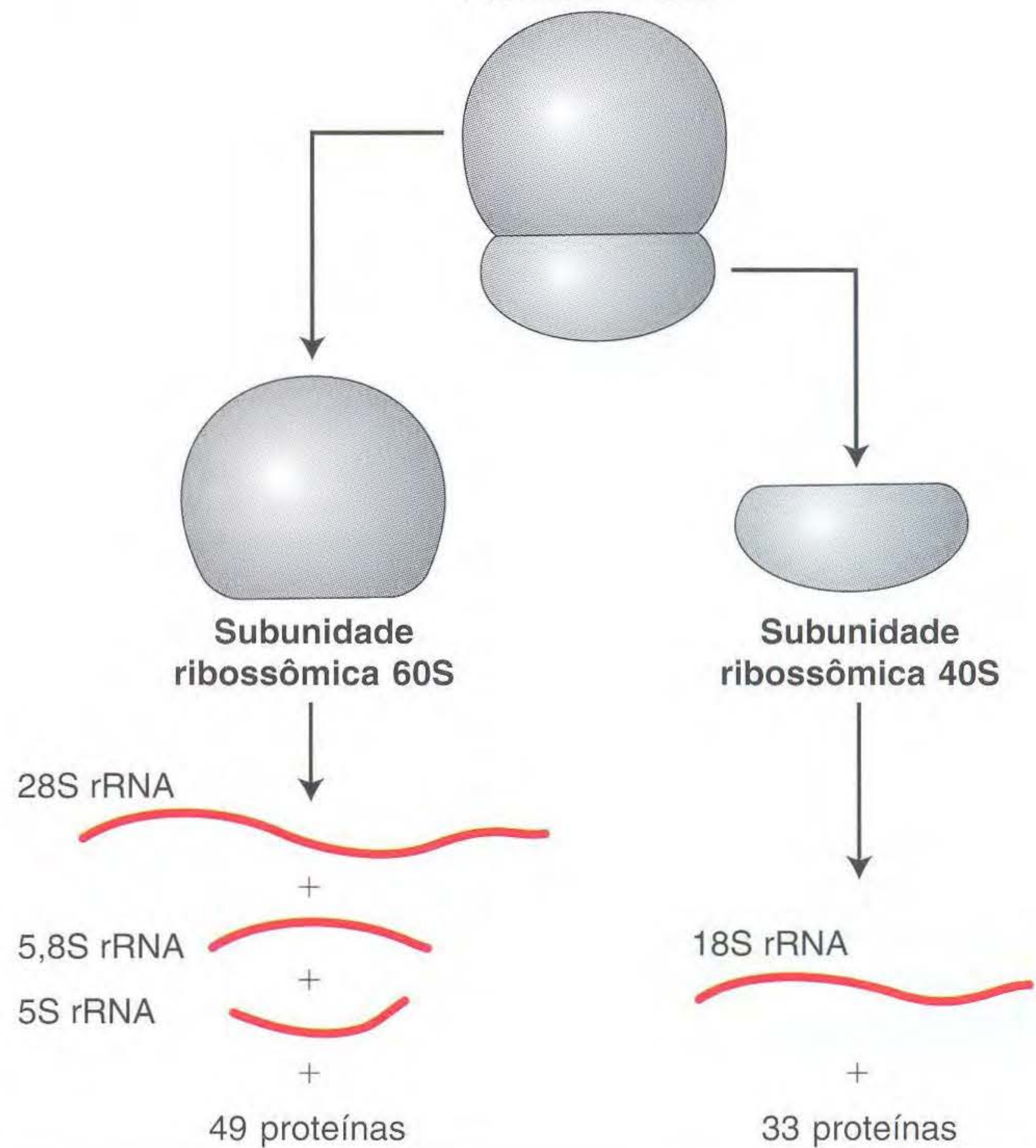


FIG. 9.11 Um ribossomo contém uma subunidade grande e outra pequena. Cada subunidade contém rRNA de tamanhos variados e um conjunto de proteínas. Existem duas principais moléculas de rRNA em todos os ribossomos. Os ribossomos procarióticos também contêm um rRNA com 120 bases de tamanho que se sedimenta em 5S, enquanto os ribossomos eucarióticos têm dois pequenos rRNA: uma molécula de 5S RNA similar à 5S procariótica e uma molécula de 5,8 S com 160 bases de comprimento.

apenas ajudada pelas proteínas ribossômicas. Em outras palavras, a subunidade maior do ribossomo funciona como uma ribozima para catalisar a formação de ligações peptídicas.

O rRNA dobra-se por pareamento intramolecular de bases

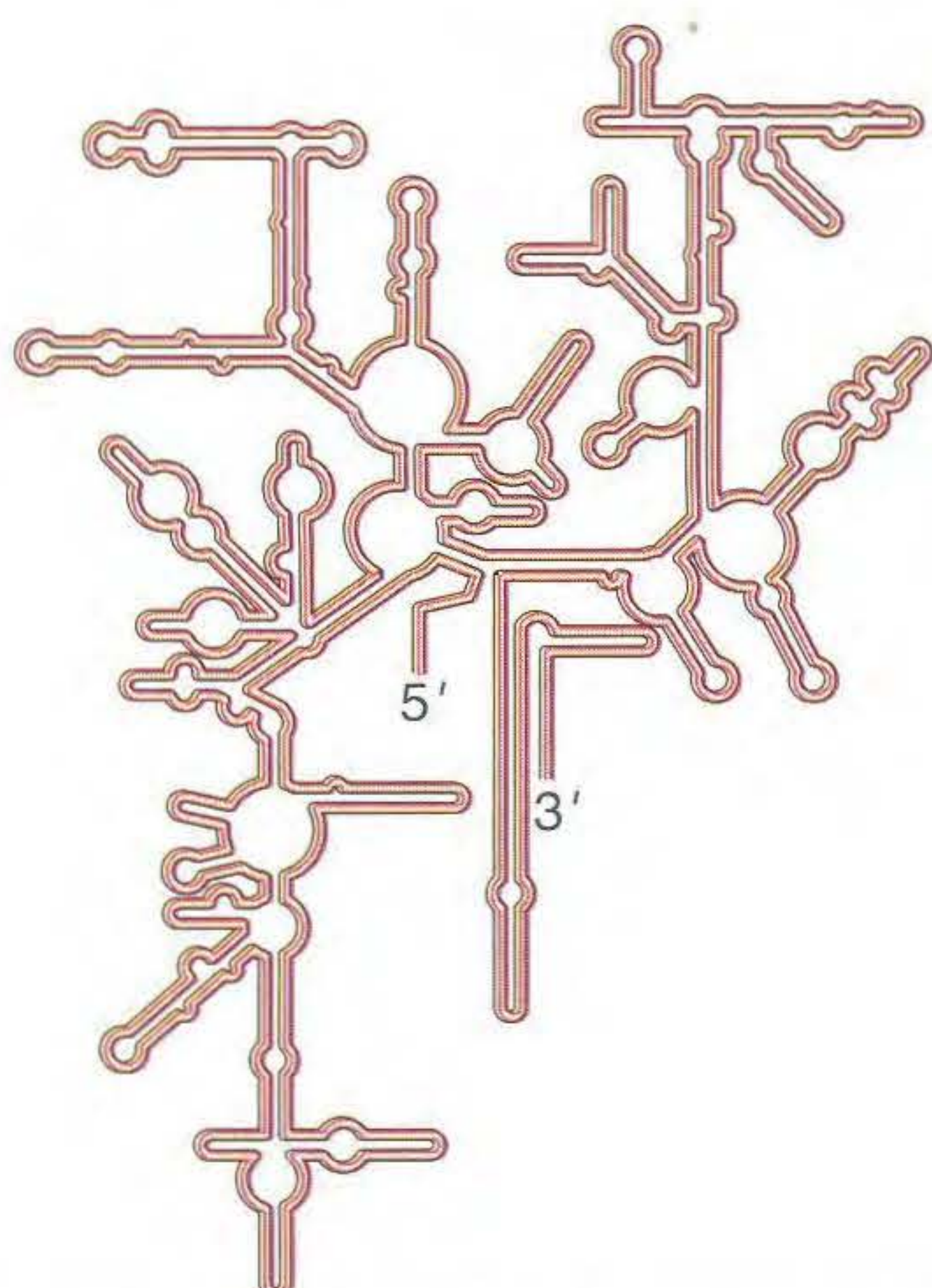


FIG. 9.12 A estrutura dobrada do RNA ribossômico 16S procariótico da pequena subunidade ribossômica.

◆ O QUE OS GENETICISTAS ESTÃO FAZENDO HOJE

Estudos estruturais similares examinaram a grande subunidade ribossômica em complexo com vários antibióticos diferentes. Esses estudos revelaram os pontos de contato entre o antibiótico e o ribossomo. Por exemplo, os macrolídios são uma família de compostos estruturalmente similares que incluem os antibióticos populares eritromicina e Zithromax. Esses antibióticos inibem a síntese de proteínas parando o ribossomo no mRNA. Eles parecem fazer isso bloqueando o chamado túnel de saída onde o polipeptídeo nascente emerge da subunidade maior (veja a Figura 9.1). Curiosamente, as bactérias patogênicas que desenvolveram resistência a alguns desses antibióticos parecem ter mutações ribossômicas que tornam o túnel de saída maior. Assim, o conhecimento de como os antibióticos se ligam aos ribossomos ajuda os cientistas a compreender como o ribossomo funciona e como projetar novos antibióticos que podem ser ativos contra mutantes resistentes. O processo do uso de informações básicas sobre a maquinaria celular para desenvolver novos antibióticos e outras drogas foi chamado de **projeto de drogas baseado na estrutura**.

Início, alongamento e término da tradução

O processo de tradução pode ser dividido em três fases: início, alongamento e término. Além do ribossomo, o mRNA,

Sítios principais de interação no ribossomo

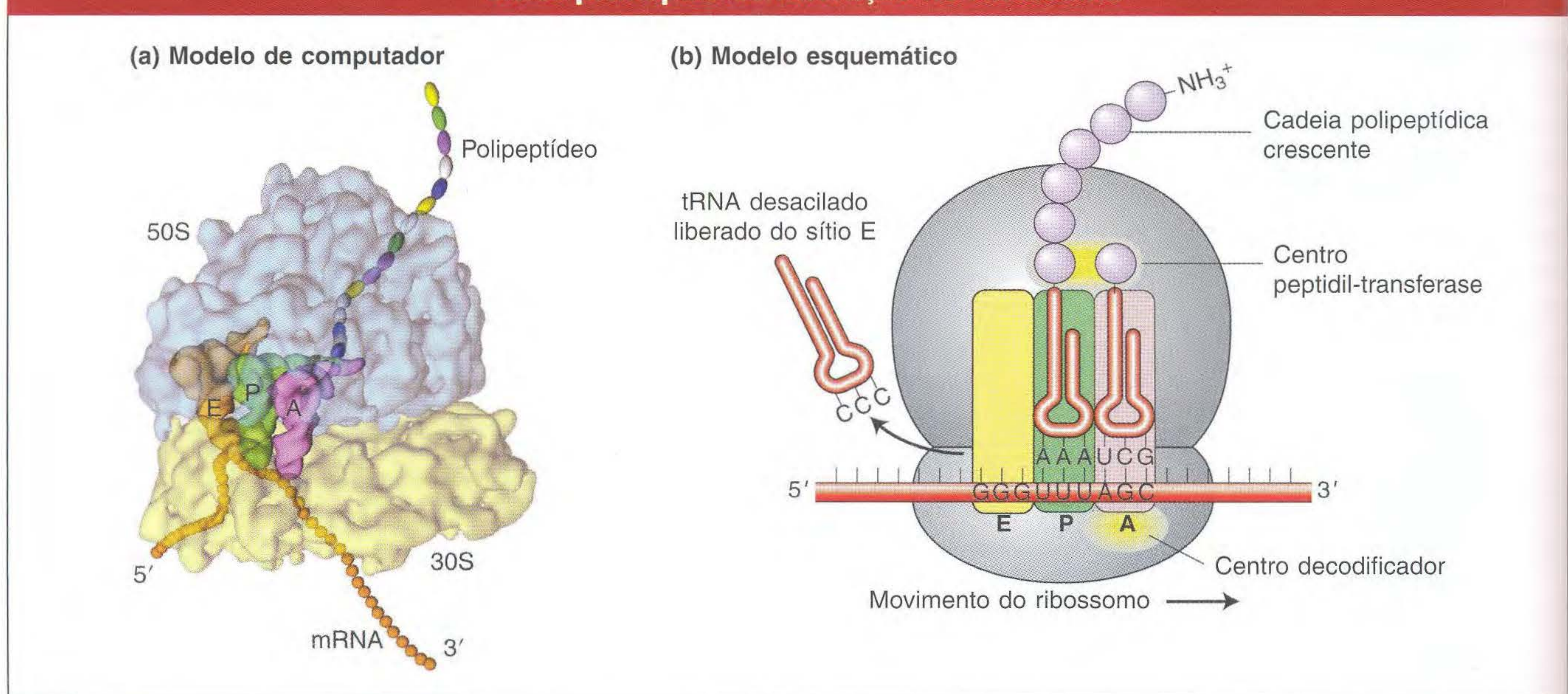


FIG. 9.13 Sítios importantes de interação em um ribossomo na fase de alongamento da tradução. (a) Um modelo de computador da estrutura tridimensional do ribossomo incluindo o mRNA, tRNA e a cadeia polipeptídica nascente à medida que emerge da subunidade ribossômica maior. (b) Um modelo esquemático do ribossomo durante o alongamento da tradução. Veja texto para detalhes. [(a) De J. Frank, *Bioessays* 23, 2001, 725-732. Fig. 2.]

tRNA e proteínas adicionais são necessários para o término bem-sucedido de cada fase. Como algumas etapas na iniciação diferem significativamente em procariontes e eucariontes, o início é descrito separadamente para os dois grupos. As fases de alongamento e término são amplamente descritas à medida que ocorrem nas bactérias, que foram o foco de muitos estudos recentes da tradução.

Início da tradução A principal tarefa da iniciação é colocar o primeiro aminoacil-tRNA no sítio P do ribossomo e, desse modo, estabelecer a correta matriz de leitura do mRNA. Na maioria dos procariontes e em todos os eucariontes, o primeiro aminoácido em qualquer polipeptídeo recém-sintetizado é a metionina, especificada pelo códon AUG. Ele é inserido não por tRNA^{Met}, mas por um tRNA especial chamado de **iniciador**, simbolizado por tRNA^{Met}_i. Em bactérias, um grupo formil é adicionado à metionina enquanto o aminoácido é ligado ao iniciador, formando a N-formilmetionina. (O grupo formil na N-formilmetionina é removido depois.)

Como a maquinaria de tradução sabe onde começar? Em outras palavras, como o códon de iniciação AUG é selecionado entre tantos códons AUG em uma molécula de mRNA? Lembre que, tanto em procariontes quanto em eucariontes, o mRNA tem uma região 5' não-traduzida que consiste na sequência entre o sítio de início da transcrição e o sítio de início traducional. Como você verá em seguida, a sequência de nucleotídeos da UTR 5' adjacente ao AUG iniciador é crítica para a ligação do ribossomo em procariontes, mas não em eucariontes.

Iniciação em procariontes Os códons de iniciação são precedidos por sequências especiais, chamadas de **sequências de Shine-**

Dalgarno, que fazem par com a ponta 3' de um rRNA, chamado rRNA 16S, na subunidade ribossômica 30S. Esse pareamento posiciona corretamente o códon iniciador no sítio P onde o tRNA iniciador irá ligar-se (Figura 9.14). O mRNA só pode parear com a subunidade 30S que está dissociada do resto do ribossomo. Note novamente que o rRNA faz uma função importante em garantir que o ribossomo esteja no lugar correto para iniciar a tradução.

Três proteínas — IF1, IF2 e IF3 (de **fator de iniciação**) — são necessárias para a correta iniciação (Figura 9.15). IF3 é necessária para manter a subunidade 30S dissociada da subunidade 50S, e para que IF1 e IF2 atuem para garantir que apenas o tRNA iniciador entre no sítio P. A subunidade 30S, o mRNA e o tRNA iniciador constituem o complexo de iniciação. O ribossomo 70S completo é formado pela associação da subunidade maior 50S com o complexo de iniciação e a liberação dos fatores de iniciação.



FIG. 9.14 Em bactérias, as bases complementares entre a ponta 3' do rRNA 16S da subunidade ribossômica pequena e a sequência de Shine-Dalgarno do mRNA posicionam o ribossomo para iniciar corretamente a tradução no códon AUG posterior.

Início da tradução em procariontes

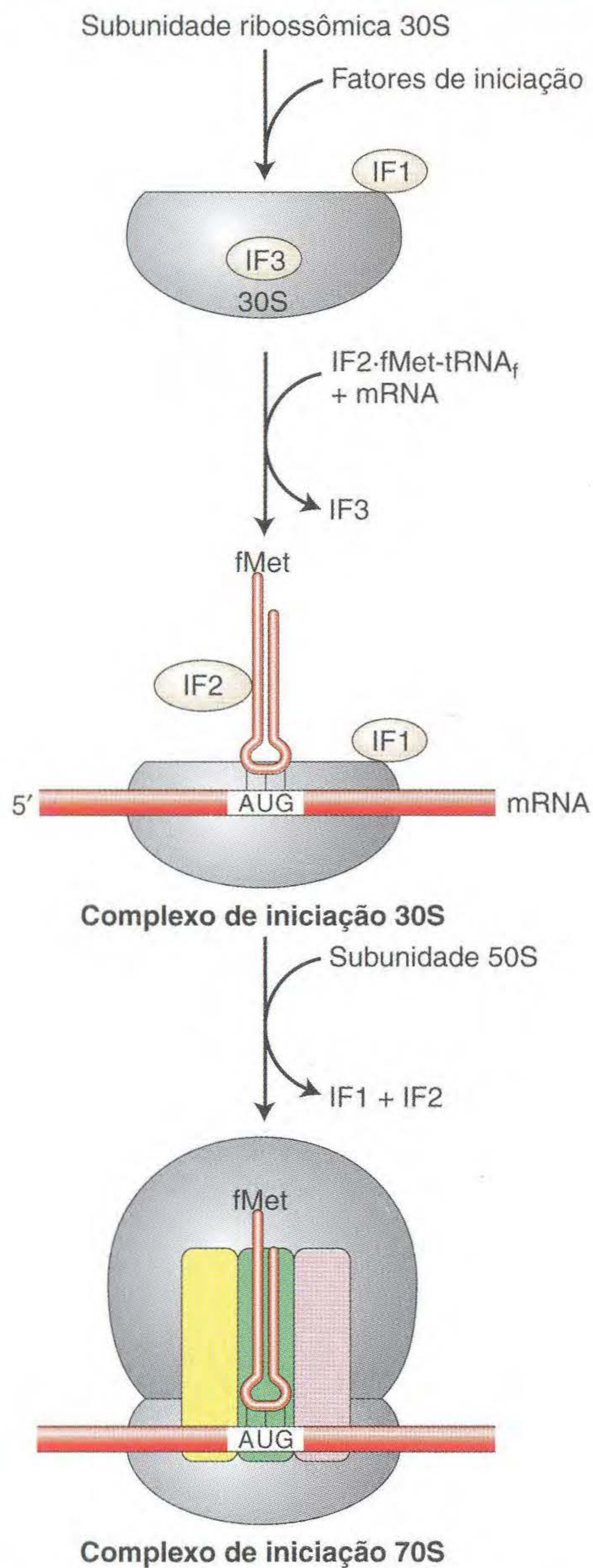


FIG. 9.15 Fatores de iniciação ajudam a montagem do ribossomo no ponto de início da tradução e, então, dissociam-se antes da tradução. [De J. Berg, J. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed. Copyright 2002 by W.H. Freeman and Company.]

Como um procarionte não tem um compartimento nuclear que separe a transcrição e a tradução, o complexo procariontário de iniciação é capaz de se formar na sequência de Shine-Dalgarno perto da ponta 5' de um RNA que ainda está sendo transcrito. Assim, a tradução pode começar nos RNA procariontários mesmo antes que eles tenham sido completamente transcritos.

Iniciação em eucariontes A transcrição e a tradução ocorrem em compartimentos separados na célula eucariótica. Como discutido no Capítulo 8, os mRNA eucarióticos são transcritos e processados no núcleo antes de serem exportados

Início da tradução em eucariontes

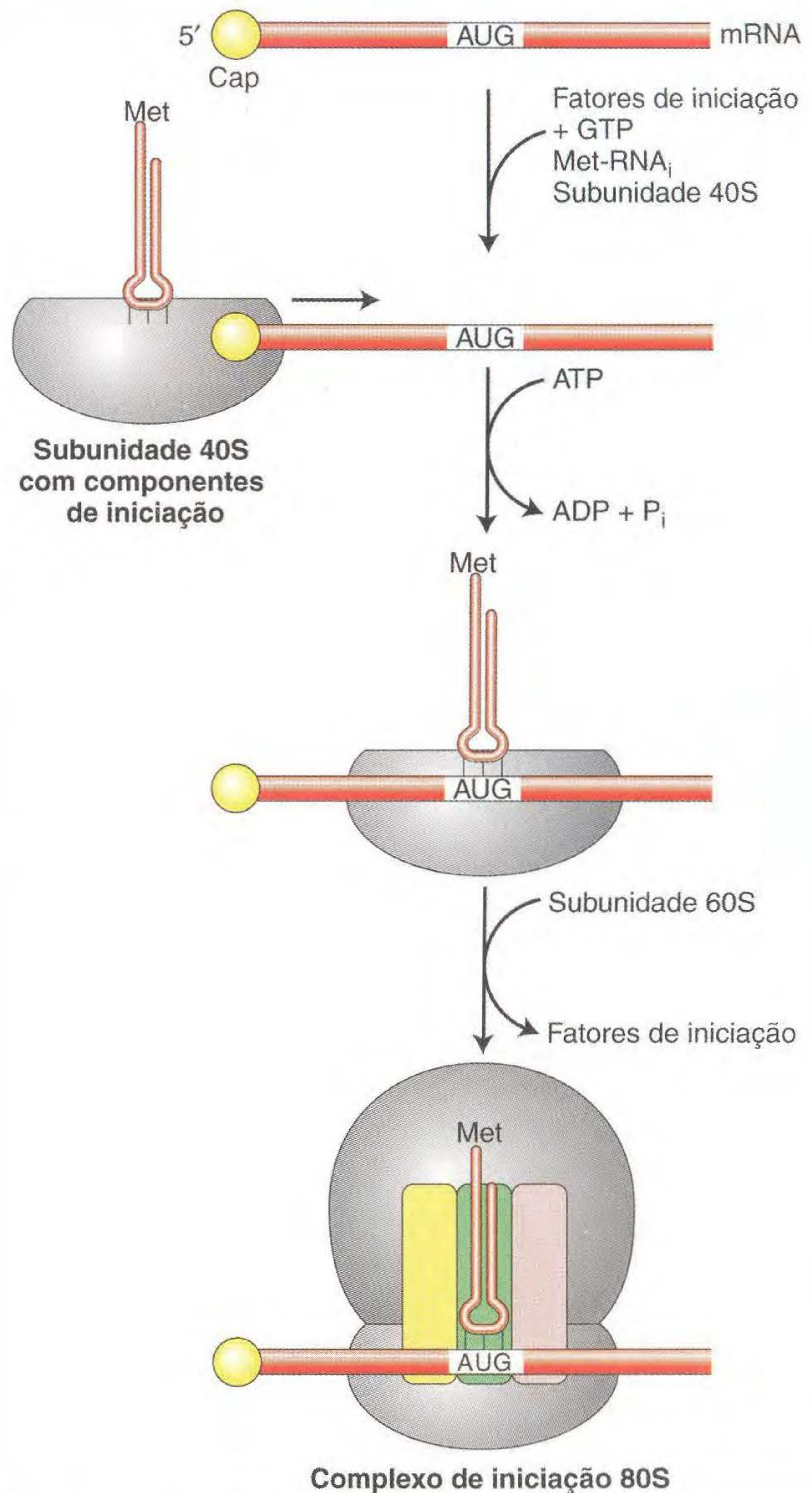


FIG. 9.16 O complexo de iniciação forma-se na ponta 5' do mRNA e o percorre no sentido de 3' à procura de um códon de início. O reconhecimento do código de início dispara a montagem do ribossomo completo e a dissociação dos fatores de iniciação (não mostrados). A hidrólise de ATP fornece energia para ativar o processo de varredura. [De J. Berg, J. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed. Copyright 2002 by W.H. Freeman and Company.]

para o citoplasma para tradução. Na chegada ao citoplasma, o mRNA é geralmente coberto de proteínas, e regiões podem ter dupla hélice devido a pareamento de base intramolecular. Essas regiões de estrutura secundária devem ser removidas para expor o códon iniciador AUG. Essa remoção é feita por fatores eucarióticos de iniciação chamados eIF4A, B e G. Esses fatores de iniciação associam-se à estrutura do cap (encontrado na ponta 5' de praticamente todos os mRNA eucarióticos) e à subunidade 40S e tRNA iniciador para formar um complexo de iniciação. Uma vez no lugar, o complexo move-se no sentido 5' para 3' e desenrola as regiões com pareamento de base (Figura 9.16). Ao mesmo tempo,

a sequência exposta é percorrida à procura de um códon AUG onde possa começar a tradução. Após o códon AUG ser apropriadamente alinhado com tRNA iniciador, o complexo de iniciação é unido à subunidade 60S para formar o ribossomo 80S. Como nos procariontes, os fatores de iniciação eucarióticos dissociam-se para formar o ribossomo antes que a fase de alongamento da tradução comece.

Alongamento É durante o processo de alongamento que o ribossomo mais se assemelha a uma fábrica. O mRNA atua como um mapa especificando a entrega de tRNA cognatos, cada um levando uma carga de aminoácido. Cada aminoácido é adicionado à cadeia polipeptídica crescente enquanto o tRNA desacilado é reciclado pela adição de outro aminoácido. A Figura 9.17 detalha as etapas do alongamento. Os fatores proteicos, chamados fatores de alongamento Tu (EF-Tu) e fator de alongamento G (EF-G), ajudam o processo de alongamento.

Como já foi descrito antes neste capítulo, um aminoacil-tRNA é formado pela ligação covalente de um aminoácido à ponta 3' de um tRNA que contém o anticódon correto. Antes que os aminoacil-tRNA possam ser usados na síntese de proteínas, eles se associam ao fator EF-Tu para formar um complexo ternário composto de tRNA, aminoácido e EF-Tu.

O ciclo de alongamento começa com um tRNA iniciador (e sua metionina associada) no sítio P e com um sítio A pronto para aceitar um complexo ternário (veja Figura 9.17). Qual dos 20 diferentes complexos ternários aceitar é determinado pelo reconhecimento códon-anticódon no centro de decodificação da subunidade menor (veja Figura 9.13b). Quando há uma correspondência correta, o ribossomo muda de forma. EF-Tu deixa o complexo ternário, e as duas pontas aminoacil são justapostas no centro peptidiltransferase da subunidade maior (veja Figura 9.13b). Nesse local é formada uma ligação peptídica com a transferência da metionina no sítio P para o aminoácido no sítio A. Nesse ponto, o segundo fator proteico, EF-G, desempenha seu papel. O fator EF-G parece se ajustar ao sítio A. Sua entrada nesse sítio muda os tRNA nos sítios A e P para os sítios P e E, respectivamente, e o mRNA move-se pelo ribossomo de modo que o códon seguinte é posicionado no sítio A (veja Figura 9.17). Quando EF-G deixa o ribossomo, o sítio A está aberto para aceitar o complexo ternário seguinte.

Nos ciclos subsequentes, o sítio A é preenchido com um novo complexo ternário conforme o tRNA desacilado deixa o sítio E. À medida que o alongamento progride, o número de aminoácidos no peptidil-tRNA (no sítio P) aumenta. Então, a ponta amino-terminal do polipeptídeo crescente emerge do túnel na subunidade 50S e sai do ribossomo.

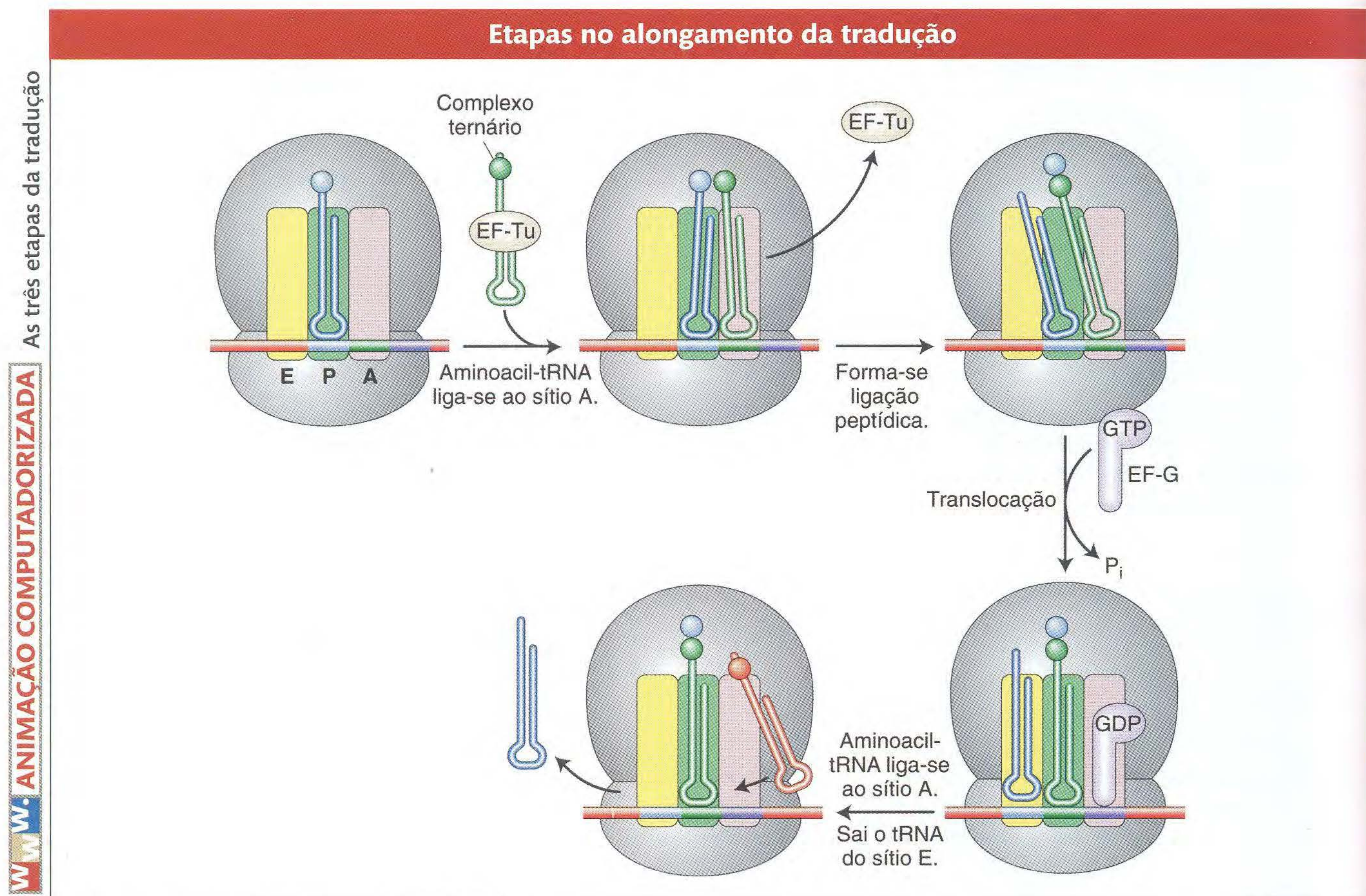


FIG. 9.17 Um complexo ternário consiste em um aminoacil-tRNA ligado a um fator EF-Tu que se liga ao sítio A. Após seu aminoácido juntar-se à cadeia polipeptídica em crescimento, um fator EF-G liga-se ao sítio A enquanto empurra os tRNA e seus códon de mRNA para os sítios E e P. Veja texto para detalhes.

Término O ciclo continua até o códon no sítio A ser um dos três códons de fim: UGA, UAA ou UAG. Lembre que nenhum tRNA reconhece esses códons. Ao contrário, proteínas chamadas **fatores de liberação** (RF1, RF2 e RF3 em bactérias) reconhecem códons de fim (Figura 9.18). Nas bactérias, RF1 reconhece UAA ou UAG, enquanto RF2 reconhece UAA ou UGA; ambos assistidos por RF3. A interação entre os fatores de liberação 1 e 2 e o sítio A difere da do complexo ternário de dois modos importantes: primeiro, os códons de fim são reconhecidos por tripeptídeos nas proteínas RF, não por um anticódon; segundo, os fatores de liberação ajustam-se ao sítio A da subunidade 30S, mas não participam na formação da ligação peptídica. Em vez disso, uma molécula de água entra no centro peptidiltransferase, e sua presença leva a uma liberação do polipeptídeo do tRNA no sítio P. As subunidades ribossômicas separam-se, e a subunidade 30S agora está pronta para formar um novo complexo de iniciação.

A tradução é feita pelos ribossomos movendo-se ao longo do mRNA no sentido 5' → 3'. Um conjunto de moléculas de tRNA leva os aminoácidos para o ribossomo, e seus anticódons ligam-se aos códons do mRNA expostos no ribossomo. Um aminoácido que chega torna-se ligado à ponta amino do polipeptídeo crescente no ribossomo.

Mutações sem sentido do supressor

É interessante considerar os supressores das mutações sem sentido definidas por Brenner *et al.* Lembre que as mutações nos fagos chamados mutantes âmbar substituem os códons tipo selvagem por códons de fim, mas as mutações de supressores no cromossomo hospedeiro contrabalançam os efeitos das mutações âmbar. Podemos agora dizer, mais especificamente, onde as mutações de supressores estavam situadas e como elas funcionam. Muitos desses supressores são mutações em genes codificantes de tRNA. Essas mutações são conhecidas por alterarem as alças do anticódon de tRNA específicos de tal modo que um tRNA torna-se capaz de reconhecer um códon de fim no mRNA. Assim, um aminoácido é inserido no códon de fim, e a tradução continua além dessa trinca. Na Figura 9.19, a mutação âmbar substitui um códon tipo selvagem pelo códon sem sentido de término da cadeia UAG. Em si, o UAG faria com que a proteína fosse prematuramente cortada na posição correspondente. A mutação supressora nesse caso produz um tRNA^{Tir} com um anticódon que reconhece o códon de fim UAG mutante. O mutante suprimido contém assim tirosina nessa posição na proteína.

O tRNA produzido por uma mutação supressora também se liga a sinais de término nas pontas das proteínas? A presença de uma mutação supressora impede assim o término normal? Muitos dos sinais naturais de término consistem em dois sinais em seguida. Devido à competição com fatores de liberação, a probabilidade de supressão em dois códons em seguida é pequena. Consequentemente, poucas cópias de proteína levam muitos aminoácidos estranhos resultantes de tradução além do códon natural de fim.

► 9.6 O Proteoma

O Capítulo 8 começou com uma discussão sobre o número de genes no genoma humano e como esse número (cerca de 25.000) era muito inferior ao número real de proteínas em uma célula humana (mais de 100.000). Agora que você sabe como a informação codificada no DNA é transcrita em RNA, e como o RNA é traduzido em proteínas, é um bom momento para revisar esse assunto e observar mais de perto as fontes de diversificação de proteínas. Primeiro, vamos rever alguns termos antigos e adicionar um novo que será útil nesta discussão. Você já sabe que o *genoma* é todo o conjunto de material genético em um conjunto cromossômico. Você aprenderá, no Capítulo 13, que o *transcriptoma* é a coleção completa de seqüências transcritas do genoma (incluindo mRNA e ncRNA). Outro termo é o **proteoma**, que foi resumidamente introduzido no Capítulo 8, mas é

Término da tradução

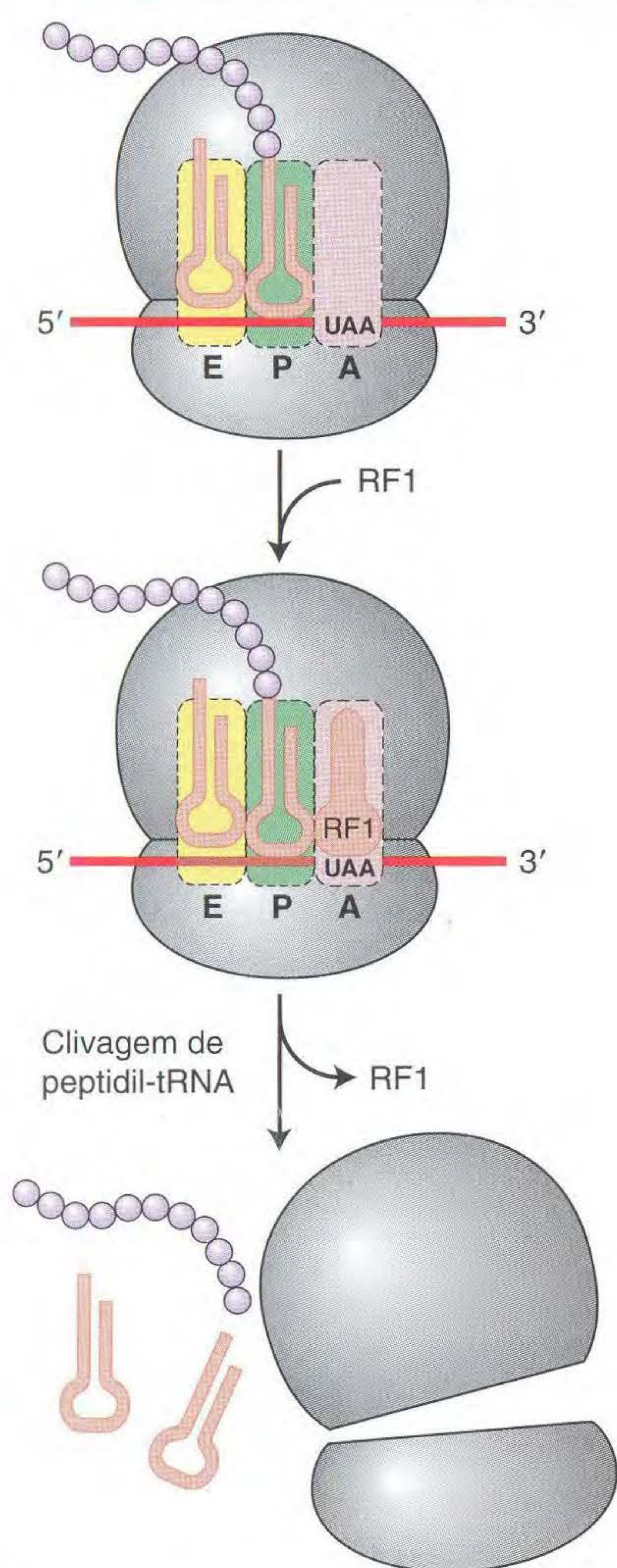
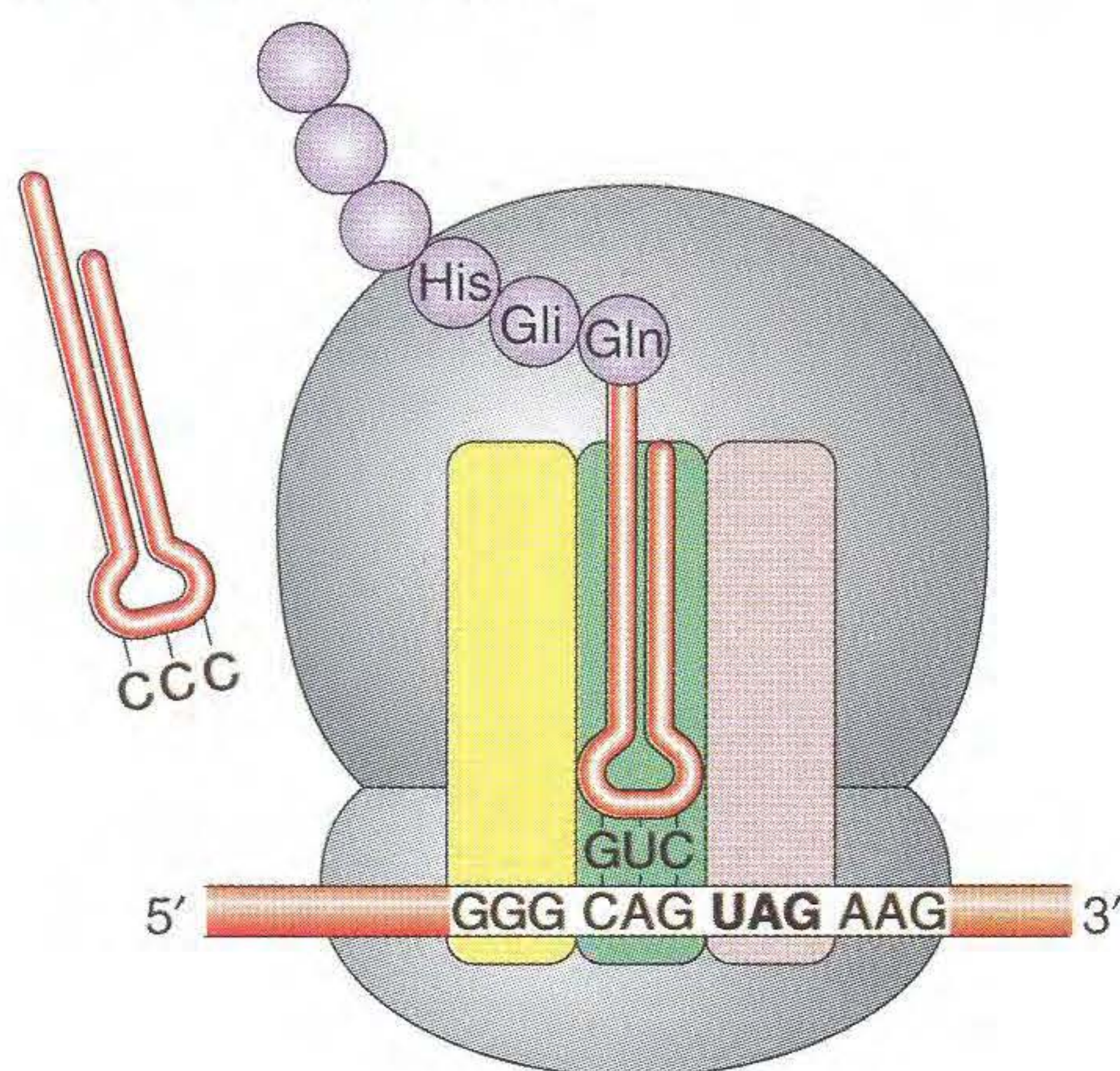


FIG. 9.18 A tradução é terminada quando os fatores de liberação reconhecem os códons de fim no sítio A do ribossomo. [De H. Lodish et al., *Molecular Cell Biology*, 5th ed. Copyright 2004 by W.H. Freeman and Company.]

Um supressor contrabalança os efeitos de uma mutação sem sentido

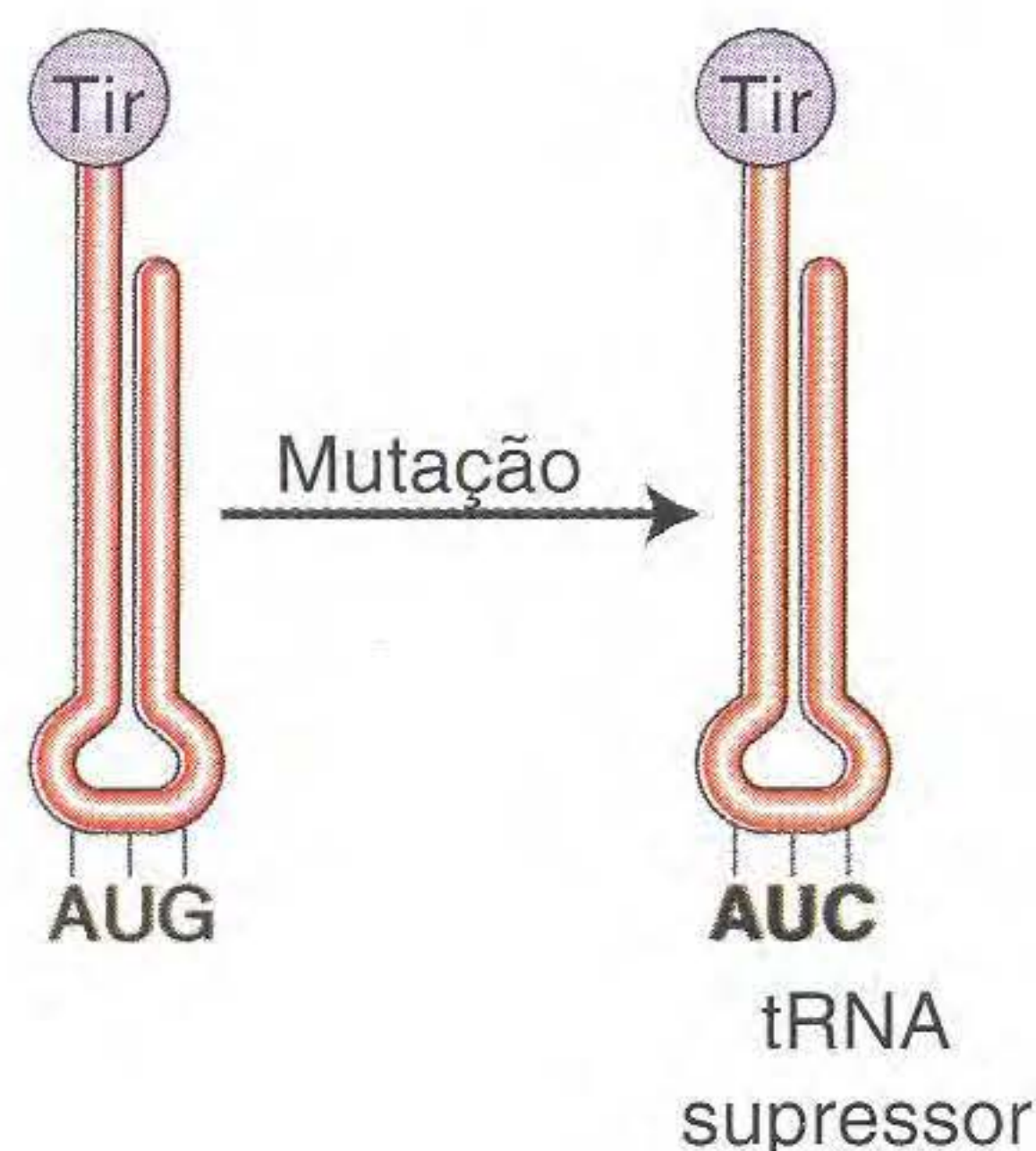
A mutação sem sentido rod^{ns}

(a) A mutação âmbar introduz um códon de fim UAG. A tradução pára.



O supressor sem sentido de tRNA

(b) O anticódon de tRNA de tirosina é mudado para AUC.



A supressão sem sentido do alelo rod^{ns}

(c) O tRNA de tirosina lê o códon UAG. A tradução continua.

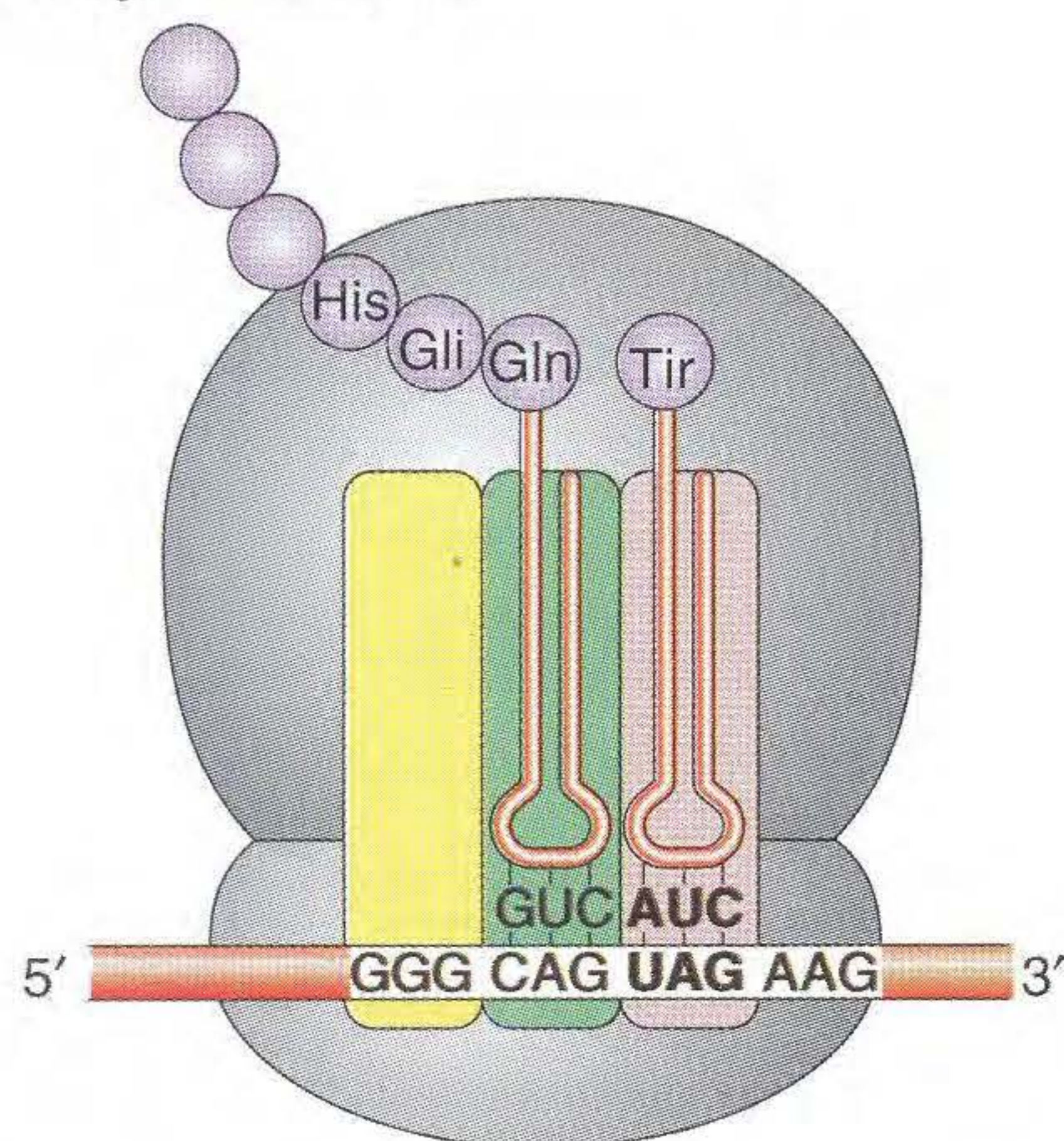


FIG. 9.19 Um supressor permite que a tradução continue quando, de outro modo, uma mutação a teria parado. (a) Término da tradução. Aqui, o aparelho de tradução não pode ir além de um códon sem sentido (UAG no caso), pois nenhum tRNA pode reconhecer a trinca UAG. Portanto, a síntese de proteínas termina, com a subsequente liberação do fragmento polipeptídico. Os fatores de liberação não são mostrados aqui. (b) As consequências moleculares de uma mutação que altera o anticódon de um tRNA de tirosina. Esse tRNA agora pode ler o códon UAG. (c) A supressão do códon UAG pelo tRNA alterado, que agora permite o alongamento da cadeia. [De D. Watson, J. Tooze, and D.T. Kurtz, *Recombinant DNA: A Short Course*. Copyright 1983 by W.H. Freeman and Company.]

aqui definido como o conjunto completo de proteínas que podem ser expressas pelo material genético de um organismo. No restante deste capítulo, você verá como o proteoma é enriquecido por dois processos celulares: a recomposição alternativa do pré-mRNA e a modificação pós-traducional das proteínas.

A recomposição alternativa gera isoformas de proteína

Como você se recorda, a recomposição alternativa do pré-mRNA permite que um gene codifique mais de uma proteína. As proteínas são feitas de domínios funcionais que são, em geral, codificados por éxons diferentes. Assim, a recomposição alternativa de um pré-mRNA pode levar à síntese de múltiplas proteínas (chamadas de **isoformas**) com combinações diferentes de domínios funcionais. Esse conceito é ilustrado por *FGFR2*, um gene humano que codifica o receptor que liga os fatores de crescimento de fibroblastos e, então, traduz um sinal dentro da célula (Figura 9.20). A proteína *FGFR2* é feita de vários domínios, inclusive um domínio extracelular ligando a proteínas. A recomposição alternativa resulta em duas isoformas que diferem em seus domínios extracelulares. Devido a essa diferença, cada isoforma liga-se a diferentes fatores de crescimento. Para muitos genes que são alternativamente recompostos, são feitas isoformas diferentes em tecidos diferentes.

Eventos pós-traducionais

Quando liberadas de um ribossomo, as proteínas recém-sintetizadas são, em sua maioria, incapazes de funcionar. Esse fato pode ser uma surpresa para os que acreditam que as seqüências de proteínas codificadas no DNA e transcritas em mRNA é tudo o que é necessário para explicar como os organismos funcionam. Como você verá nesta seção e nos capítulos subsequentes deste livro, a seqüência de DNA é apenas parte da história. Nesse caso, todas as proteínas recém-sintetizadas necessárias para se dobrar corretamente e os aminoácidos de algumas proteínas precisam ser quimicamente modificados. Como o dobramento da proteína e a modificação ocorrem após a síntese de proteínas, eles são chamados de eventos pós-traducionais.

Dobramento da proteína dentro da célula O mais importante evento pós-traducional é o dobramento da proteína nascente (recém-sintetizada) em sua correta forma tridimensional. Uma proteína que é dobrada corretamente é dita estando em sua conformação nativa (em contraste com uma proteína não-dobrada ou mal dobrada, que é inativa). Como vimos no começo deste capítulo, as proteínas existem em uma incrível diversidade de estruturas. As estruturas distintas das proteínas são essenciais para sua atividade enzimática, para sua habilidade em se ligar ao DNA e por desempenhar papéis estruturais na célula. Embora se saiba, desde os anos de 1950, que a seqüência de aminoácidos de uma proteína determina sua estrutura tridimensional, também é verdade que o ambiente aquoso dentro da célula não favorece o dobramento correto da maioria das proteínas.

A recomposição alternativa produz isoformas proteicas correlatas mas distintas

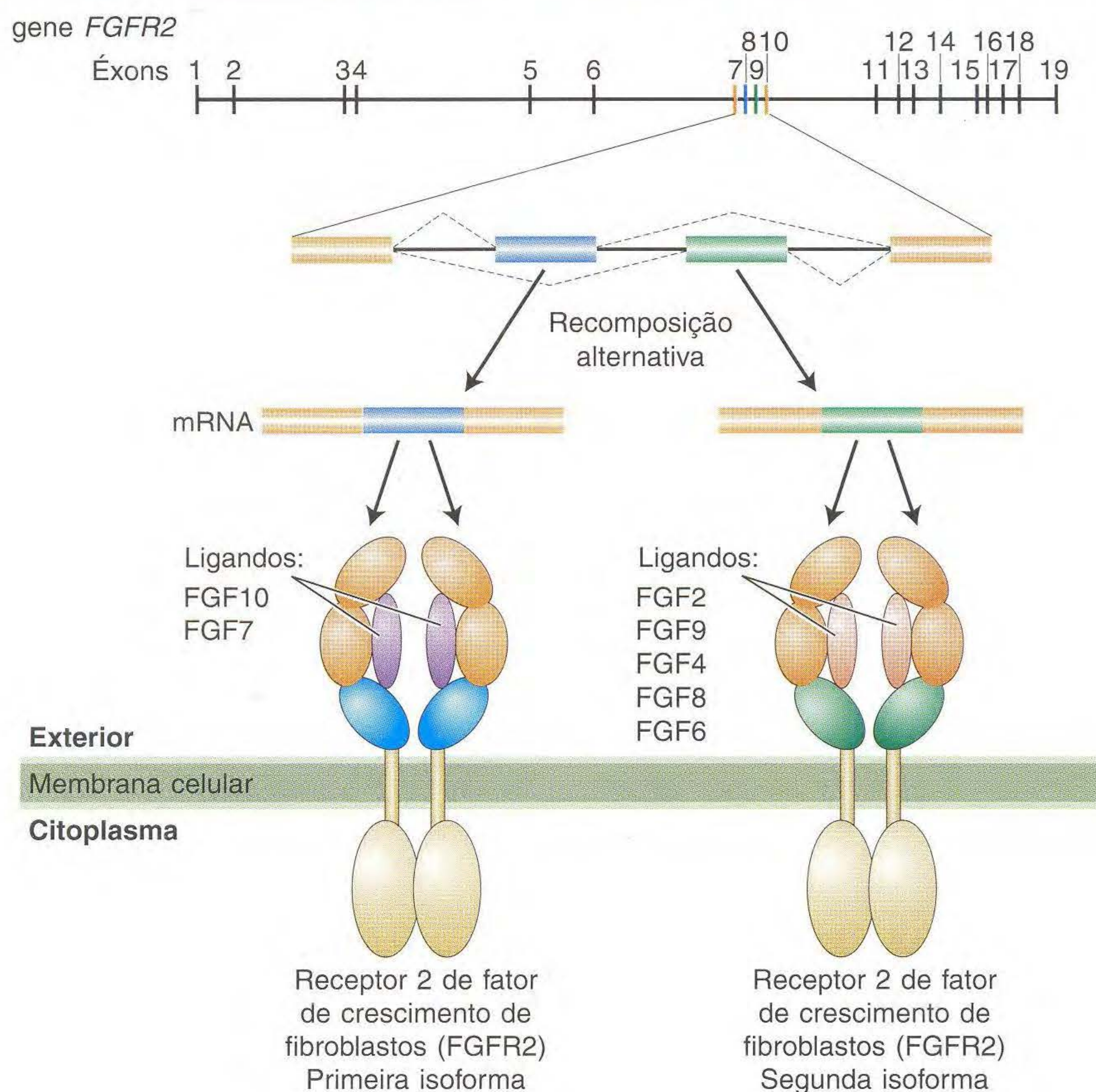


FIG. 9.20 RNA mensageiros produzidos por recomposição alternativa do pré-mRNA do gene humano *FGFR2* codificam duas isoformas de proteína que se unem a ligandos diferentes (os fatores de crescimento).

Considerando que as proteínas dobram-se corretamente na célula, há uma dúvida de longa duração: como é feito esse dobramento correto?

A resposta parece ser que as proteínas nascentes são dobradas corretamente com a ajuda de chaperonas, uma classe de proteínas encontradas em todos os organismos, desde bactérias até plantas e humanos. Uma família de chaperonas, chamada de chaperoninas GroE, forma grandes complexos multissubunitários chamados máquinas de dobramento chaperonina. Embora o mecanismo exato não seja compreendido, as proteínas não-dobradas recém-sintetizadas são consideradas como entrando em uma câmara na máquina de dobramento que fornece um microambiente eletricamente neutro dentro do qual a proteína nascente pode, bem-sucedidamente, dobrar-se em sua conformação nativa.

Modificação pós-traducional de cadeias laterais de aminoácidos Como já foi visto, as proteínas são polímeros de aminoácidos feitos de qualquer um dos 20 tipos diferentes. Entretanto, a análise bioquímica de muitas proteínas revela que uma variedade de moléculas podem ser ligadas covalentemente às cadeias laterais de aminoácidos. Mais de 300

modificações de cadeias laterais de aminoácidos podem ocorrer após a tradução. Duas das modificações pós-traducionais mais comumente encontradas — fosforilação e ubiquitinização — são consideradas em seguida.

Fosforilação As enzimas chamadas de *cinases* ligam grupos fosfato aos grupos hidroxila dos aminoácidos serina, treonina e tirosina, enquanto as enzimas chamadas de *fosfatases* removem esses grupos fosfato. Como os grupos fosfato têm carga negativa, sua adição a uma proteína geralmente muda a conformação da proteína. A adição/remoção de grupos fosfato serve como um interruptor reversível para controlar uma variedade de eventos celulares, incluindo a atividade enzimática, interações proteína-proteína e interações proteína-DNA (Figura 9.21).

Uma medida da importância da fosforilação de proteínas é o número de genes que codificam a atividade de cinases no genoma. Mesmo um organismo simples, tal como uma levedura, tem centenas de genes de cinases, enquanto a planta *Arabidopsis thaliana* tem mais de 1.000. Outra medida do significado da fosforilação de proteínas é que a maioria das numerosas interações proteína-proteína que ocorrem em

Fosforilação e desfosforilação de proteínas

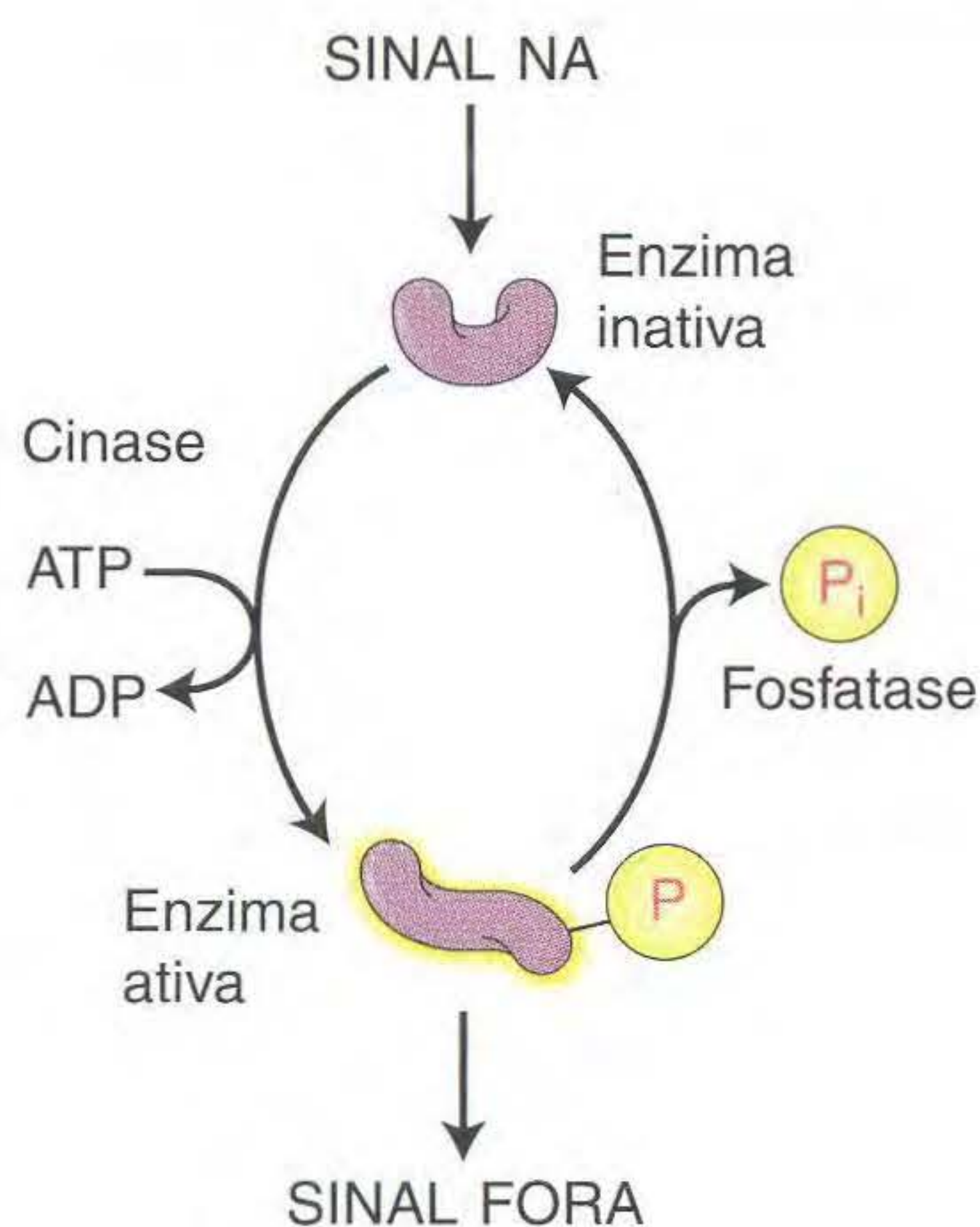


FIG. 9.21 As proteínas podem ser ativadas pela ligação enzimática de grupos fosfato aos grupos laterais de seus aminoácidos e inativar a remoção desses grupos fosfato.

uma célula típica são reguladas por fosforilação. As análises recentes das interações proteína-proteína do proteoma indicam que a maioria das proteínas funcionam interagindo com outras proteínas. **Interatoma** é o nome dado ao conjunto completo de interações de proteína. Um modo de represen-

tar a rede de interações proteína-proteína que constitui um interatoma é mostrado na Figura 9.22. Como você pode ver, o interatoma do verme nematelminto *C. elegans* é composto de centenas, talvez milhares, de interações proteína-proteína. Entretanto, o número de interações representadas na Figura 9.22 constitui apenas uma pequena fração das interações proteína-proteína que estão ocorrendo em todas as células de todos os organismos. Qual o significado biológico dessas interações? Neste capítulo e nos precedentes, você viu que as interações proteína-proteína são centrais para a função de grandes máquinas biológicas tais como o replissomo, o spliceossomo e o ribossomo.

Ubiquitinização Surpreendentemente, uma das mais comuns modificações pós-traducionais não é sutil, como na adição de um grupo fosfato. A adição de cadeias de várias cópias de uma proteína chamada **ubiquitina** à ϵ -amina das lisinas (chamada de **ubiquitinização**) marca a proteína para degradação por uma protease chamada de proteassomo 26S (Figura 9.23). A ubiquitina contém 76 aminoácidos e é encontrada apenas nos eucariontes, onde é altamente conservada nas plantas e animais. Duas grandes classes de proteínas são alvo para destruição pela ubiquitinização: proteínas de vida curta, tais como as reguladoras do ciclo celular, ou proteínas que se tornam danificadas ou mutadas.

Todas as interações de proteínas em um organismo compõem o interatoma

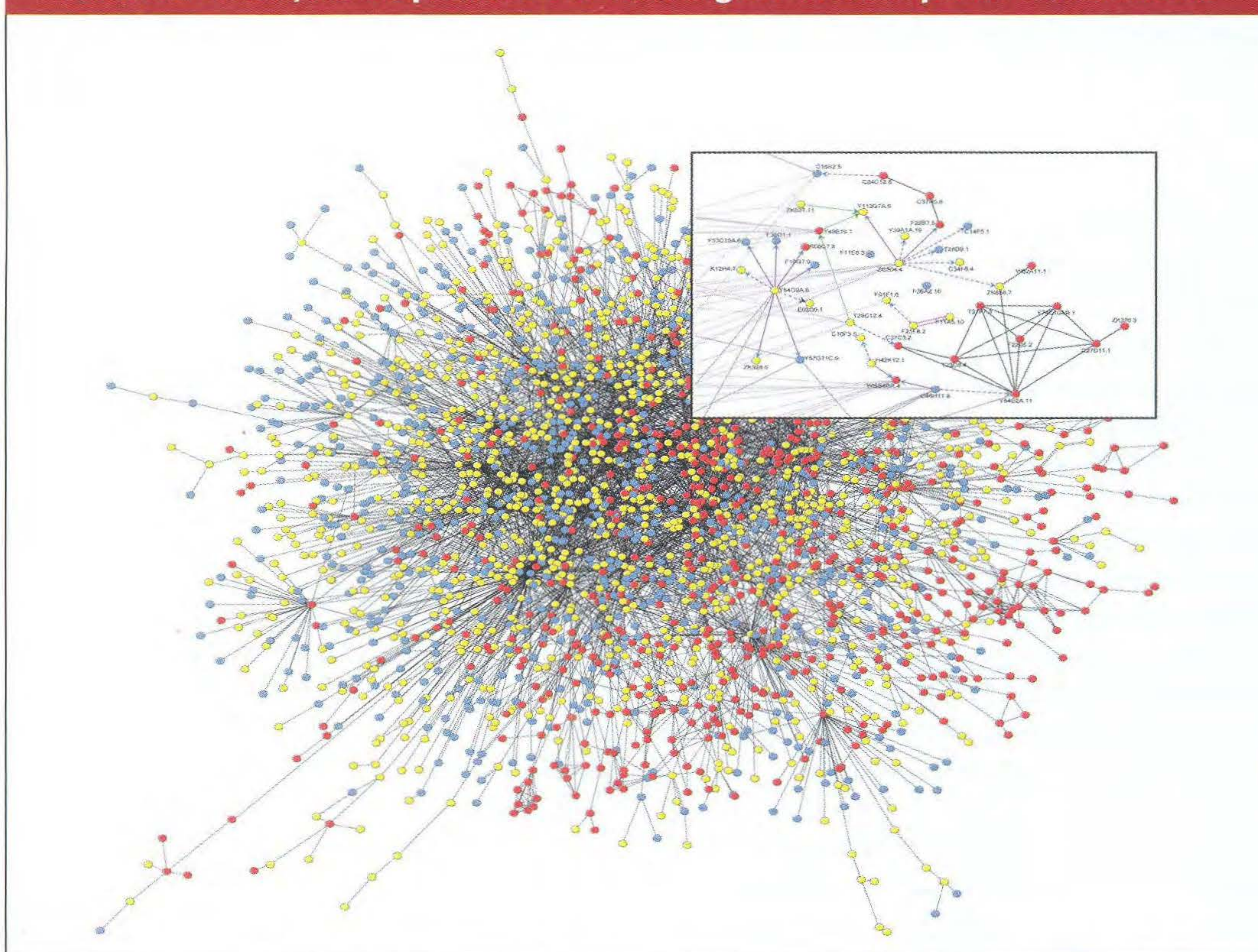


FIG. 9.22 As proteínas (representadas por círculos) interagem com outras proteínas (conectadas por linhas) para formar simples ou grandes complexos proteicos. Esse interatoma é de *C. elegans*. [Adaptado de S. Li et al., "A Map of the Interactome Network of the Metazoan *C. elegans*," *Science* 303, 2004, 540–543.]

Ubiquitinização marca uma proteína para degradação

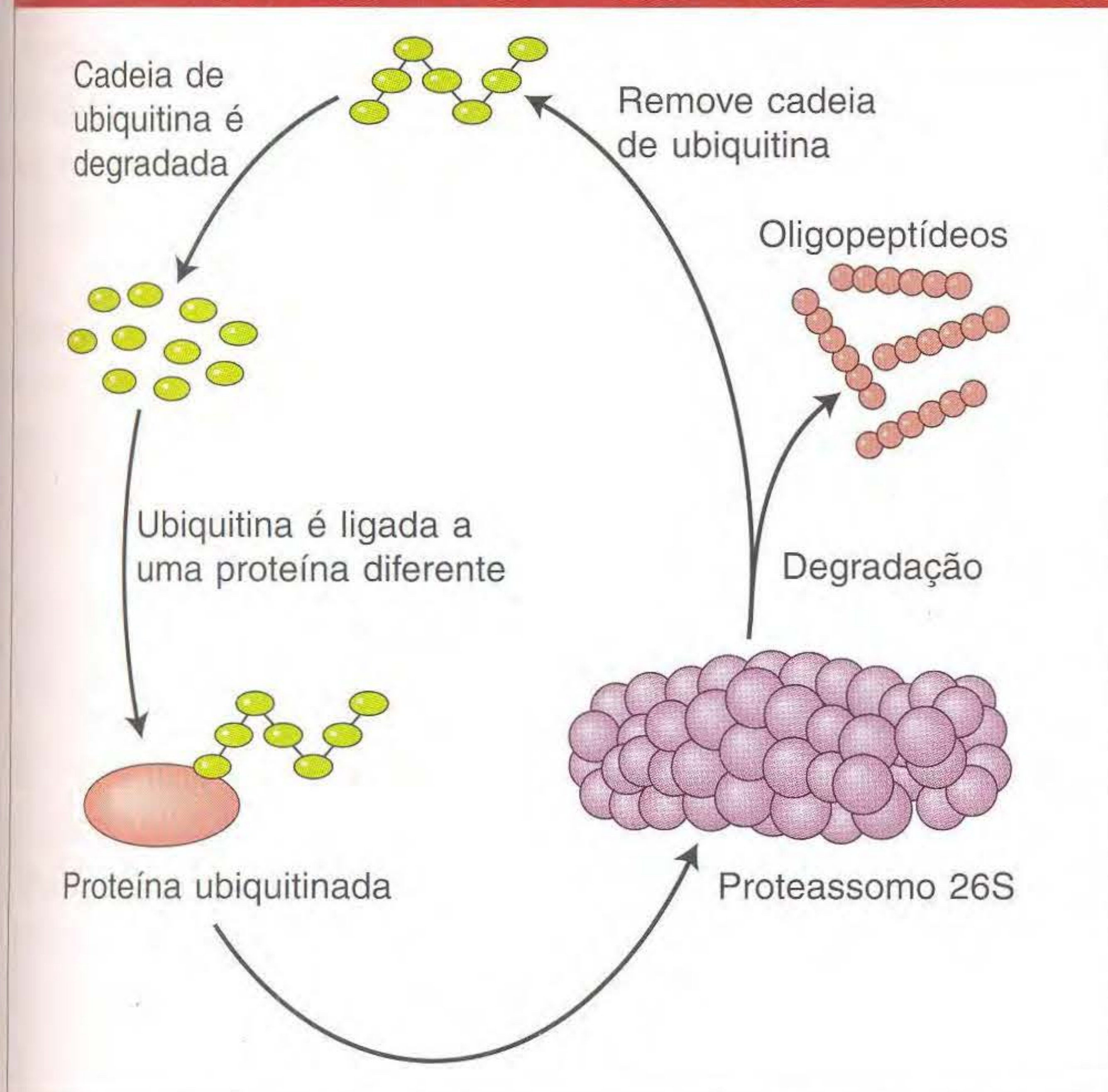


FIG. 9.23 São mostradas as principais etapas na degradação de proteínas mediada pela ubiquitina. A ubiquitina primeiro é conjugada a outra proteína e, então, degradada pelo proteassomo. A ubiquitina e os oligopeptídeos são então reciclados.

Direcionamento de proteínas Em eucariontes, todas as proteínas são sintetizadas nos ribossomos no citoplasma. Entretanto, algumas dessas proteínas terminam no núcleo, outras nas mitocôndrias, ainda outras ancoradas na membrana ou secretadas pela célula. Como essas proteínas “sabem” para onde devem ir? A resposta a esse problema aparentemente complexo na verdade é bem simples: uma proteína recém-sintetizada contém uma curta sequência que direciona a proteína para o lugar correto ou compartimento celular. Por exemplo, uma proteína da membrana recém-sintetizada ou uma proteína destinada a uma organela tem um curto peptídeo líder, chamado **sequência de sinal**, em sua ponta aminoterminal. Para proteínas da membrana, esse trecho de 15 a 25 aminoácidos direciona a proteína para canais na membrana do retículo endoplasmático onde a sequência de sinal é clivada por uma peptidase (Figura 9.24). Do retículo endoplasmático, a proteína é direcionada para seu destino final. Um fenômeno similar existe para algumas proteínas bacterianas que são secretadas.

As proteínas destinadas ao núcleo incluem a RNA e a DNA polimerases e os fatores de transcrição discutidos nos Capítulos 7 e 8. As sequências de aminoácidos inseridas nos interiores de proteínas ligadas ao núcleo são necessárias para o

Sequências de sinal marcam proteínas para secreção

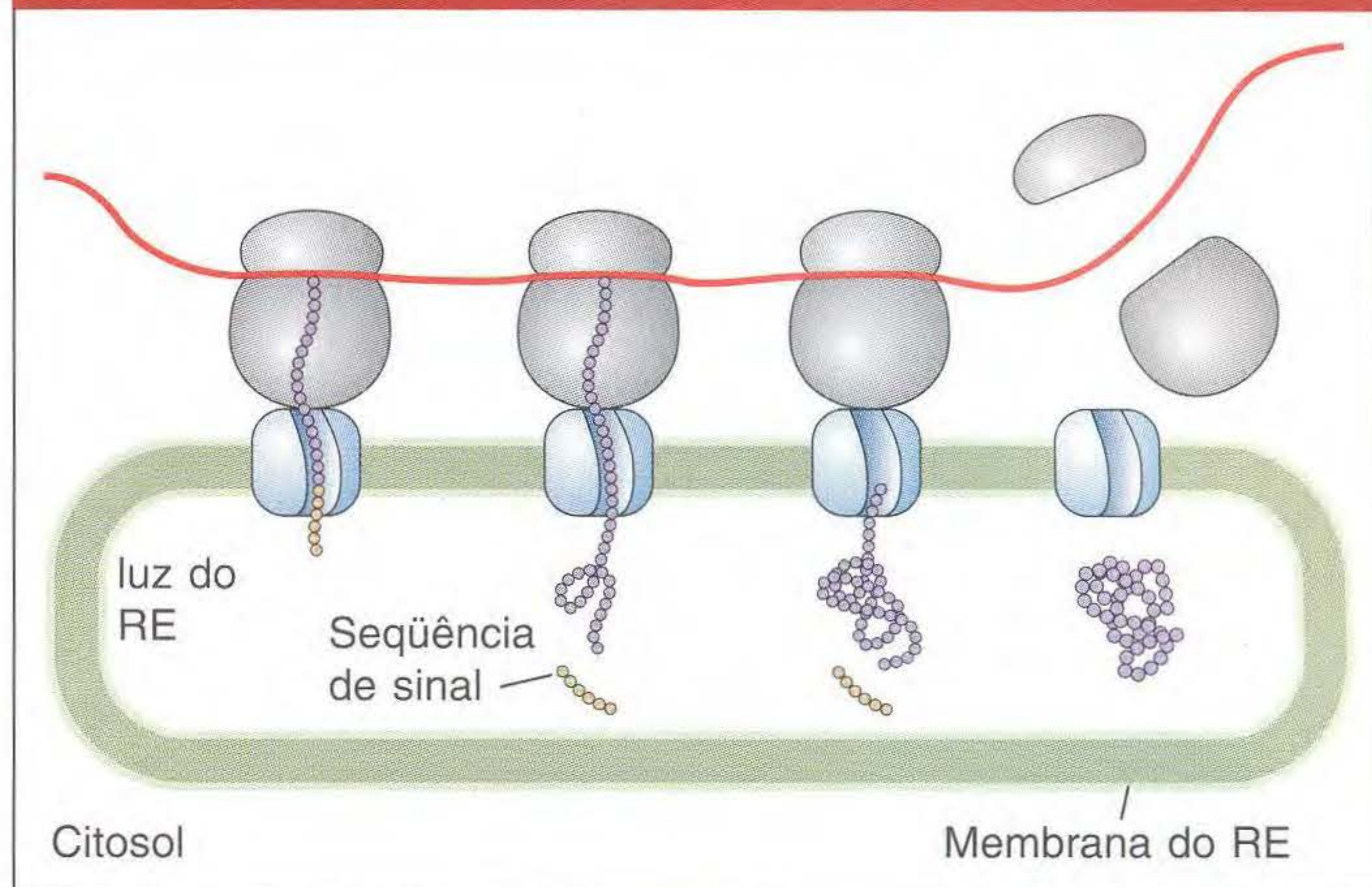


FIG. 9.24 As proteínas destinadas a serem secretadas pela célula têm uma sequência amino-terminal que é rica em unidades hidrofóbicas. Essa sequência de sinal liga-se a proteínas na membrana do retículo endoplasmático (RE) que levam o restante da proteína através da bicamada de lipídios. A sequência de sinal é clivada da proteína nesse processo por uma enzima chamada de *peptidase de sinal* (não mostrada). Uma vez dentro do retículo endoplasmático, a proteína é direcionada para a membrana celular por onde será secretada.

transporte do citoplasma para o núcleo. Essas **sequências de localização nuclear (NLS)** são reconhecidas por proteínas receptoras citoplasmáticas que transportam proteínas recém-sintetizadas pelos poros nucleares, sítios na membrana pelos quais grandes moléculas são capazes de passar para dentro e para fora do núcleo. Uma proteína não normalmente encontrada no núcleo será direcionada para o núcleo se uma NLS estiver ligada a ela.

Por que as sequências de sinal são clivadas durante o direcionamento, enquanto uma NLS, situada no interior de uma proteína, permanece após a proteína mover-se para o núcleo? Uma explicação pode ser que, na desintegração nuclear que acompanha a mitose (veja o Capítulo 2), as proteínas situadas no núcleo podem ser encontradas no citoplasma. Como tal proteína contém uma NLS, ela pode ser realocada no núcleo de uma célula-filha que resulta da mitose.

Em sua maioria, as proteínas eucarióticas são inativas, a menos que modificadas após a tradução. Alguns eventos pós-traducionais, tais como a fosforilação ou ubiquitinização, modificam grupos laterais de aminoácidos, promovendo assim ativação ou degradação, respectivamente. Outros mecanismos pós-traducionais reconhecem determinados aminoácidos em uma sequência de proteínas e direcionam essas proteínas para locais onde sua atividade é necessária dentro e fora da célula.

Resumo

Este capítulo lidou com a tradução da informação codificada na sequência de nucleotídeos de um mRNA na sequência de aminoácidos de uma proteína. Nossas proteínas, mais que

qualquer outra macromolécula, determinam quem somos e o que somos. Elas são enzimas responsáveis pelo metabolismo celular, incluindo a síntese de DNA e RNA, e são os fatores

reguladores necessários para a expressão do programa genético. A versatilidade das proteínas como moléculas biológicas é manifesta na diversidade de formas que elas podem adotar. Além disso, mesmo depois de sintetizadas, elas podem ser modificadas de uma variedade de modos pela adição de moléculas que podem alterar seu funcionamento.

Considerando o papel central das proteínas na vida, não é surpreendente que tanto o código genético quanto a maquinaria para traduzir esse código em proteínas tenham sido altamente conservados das bactérias aos humanos. Os principais componentes da tradução são as três classes de RNA: tRNA, mRNA e rRNA. A precisão da tradução depende da ligação enzimática de um aminoácido com seu tRNA cognato, gerando uma molécula de tRNA carregada. Como adaptadores, os tRNA são as moléculas-chave na tradução. Em contraste, o grande ribossomo é a fábrica onde o mRNA, tRNA carregados e outros fatores proteicos juntam-se para a síntese de proteínas.

A principal decisão na tradução é onde iniciá-la. Nos procariontes, o complexo de iniciação é montado no mRNA na sequência de Shine-Dalgarno, logo antecedente ao códon de início AUG. O complexo de iniciação em eucariontes é montado na estrutura cap 5' do mRNA e move-se na direção 3' até

que o códon de início seja reconhecido. A fase mais longa da tradução é o ciclo de alongamento. Nessa fase, o ribossomo move-se ao longo do mRNA, revelando o códon seguinte que irá interagir com seu tRNA carregado cognato, de modo que os tRNA carregados com aminoácidos podem ser adicionados à cadeia polipeptídica crescente. Esse ciclo continua até que seja encontrado um códon de fim. Os fatores de liberação facilitam o término da tradução.

Nos últimos anos, novas técnicas de imagens revelaram interações ribossômicas ao nível atômico. Com esses novos “olhos”, nós agora podemos ver que o ribossomo é uma máquina dinâmica incrível que muda de forma em resposta aos contatos feitos com tRNA e com proteínas. Além disso, as imagens a nível de resolução atômica revelaram que os RNA ribossômicos, e não as proteínas ribossômicas, estão intimamente associados aos centros funcionais do ribossomo.

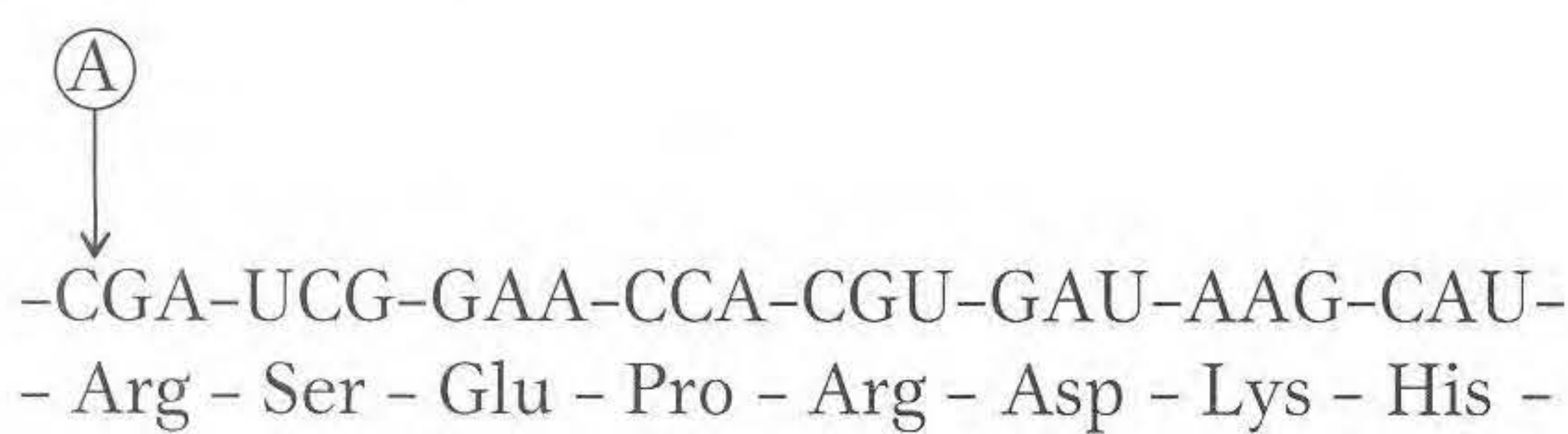
O proteoma é o conjunto completo de proteínas que pode ser expresso pelo material genético de um organismo. Enquanto um eucarionte multicelular típico tem cerca de 20.000 genes, o proteoma típico é provavelmente 10 a 50 vezes maior. Essa diferença é, em parte, o resultado de modificações pós-traducionais, tais como a fosforilação e ubiquitinação, que influenciam a atividade e estabilidade da proteína.

► Termos-chave

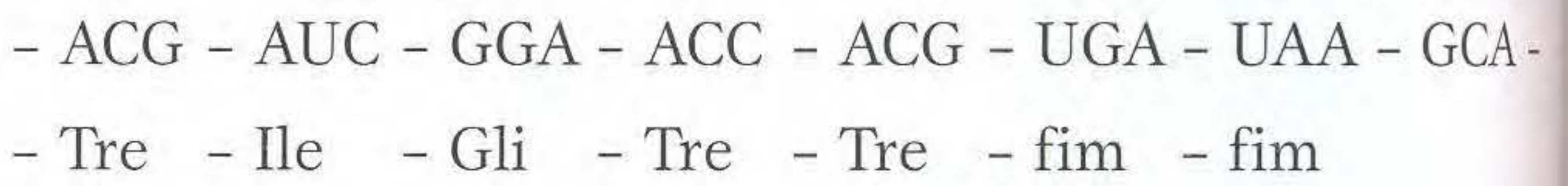
aminoácido	iniciador	sequência de localização nuclear (NLS)
aminoacil-tRNA sintetase	interatoma	sequência de Shine-Dalgarno
anticódon	isoforma	sequência de sinal
centro decodificador	oscilação	sítio A
centro peptidiltransferase	polipeptídeo	sítio ativo
código redundante	ponta amino	sítio E
códon	ponta carboxila	sítio P
colinearidade	projeto de drogas baseado na estrutura	subunidade
domínio	proteína fibrosa	trinca
estrutura primária	proteína globular	tRNA carregado
estrutura quaternária	proteoma	tRNA isoceptor
estrutura secundária	ribossomo	ubiquitina
estrutura terciária	RNA ribossômico (rRNA)	ubiquitinação
fator de iniciação	RNA transportador (tRNA)	
fator de liberação (RF)		

► Problemas Resolvidos

Problema resolvido 1. Usando a Figura 9.6, mostre as consequências na subsequente tradução da adição de uma adenina ao começo da seguinte sequência codificante:



SOLUÇÃO
Com a adição de A no começo da sequência codificante, a matriz de leitura muda, e um conjunto diferente de aminoácidos é especificado pela sequência, como mostrado aqui (note que é encontrado um conjunto de códons sem sentido, o que resulta no término da cadeia):



Problema resolvido 2. A adição de um único nucleotídeo seguida de uma deleção de um único nucleotídeo distante aproximadamente 20 pb no DNA causa uma mudança na seqüência de proteínas de

-His-Tre-Glu-Asp-Trp-Leu-His-Gln-Asp-

para

-His-Asp-Arg-Gli-Leu-Ala-Tre-Ser-Asp-

Que nucleotídeo foi adicionado e que nucleotídeo foi deletado? Quais são as seqüências de mRNA original e a nova? (**Dica:** Consulte a Figura 9.6.)

SOLUÇÃO

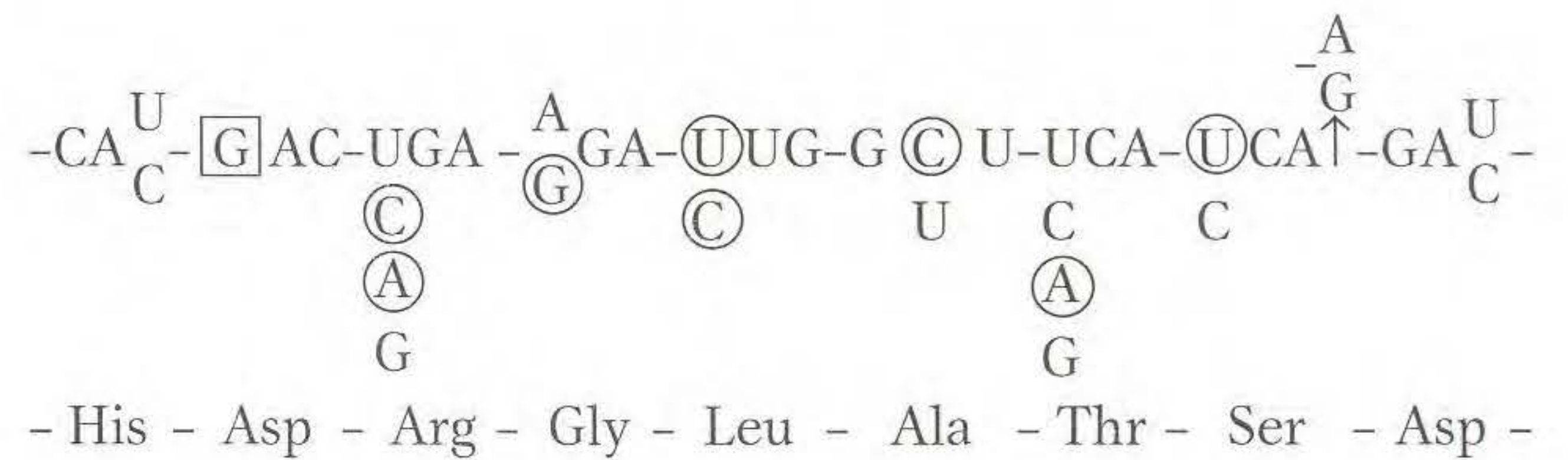
Podemos desenhar uma sequência de mRNA da proteína original (com as inerentes ambigüidades desse estágio):

-His-Thr-Glu-Asp-Trp-Leu-His-Gln-Asp

$$\begin{array}{ccccccc} & \text{U} & & & \text{U} & & \\ -\text{CA}_{\text{C}}^{\text{U}} & -\text{ACC}-\text{GA}_{\text{G}}^{\text{A}} & -\text{GA}_{\text{C}}^{\text{U}} & -\text{UGG}-\text{CUC}-\text{CA}_{\text{C}}^{\text{U}} & -\text{CA}_{\text{G}}^{\text{A}} & -\text{GA}_{\text{C}}^{\text{U}} & \\ & \text{A} & & & \text{A} & & \\ & \text{G} & & & \text{G} & & \\ & & & & \text{UUA} & & \\ & & & & \text{G} & & \end{array}$$

Como a mudança na seqüência de proteínas que nos foi dada no começo do problema começa após o primeiro aminoácido

(His) devido à adição de um único nucleotídeo, podemos deduzir que o códon de Tre deve ter mudado para um códon de Asp. Essa mudança deve resultar da adição de uma G diretamente antes do códon Tre (indicado por um box), que muda a matriz de leitura, como mostrado aqui:



Além disso, como uma deleção de um nucleotídeo deve restaurar o códon de Asp final para a correta matriz de leitura, uma A ou G deve ter sido deletada do final do códon próximo ao último original, como mostrado por uma seta. A sequência original da proteína nos permite desenhar o mRNA com um número de ambigüidades. Entretanto, a sequência de proteínas resultante da mudança de matriz de leitura nos permite determinar que nucleotídeo estava no mRNA original na maioria desses pontos de ambigüidade. O nucleotídeo que deve ter aparecido na sequência original está circulado. Apenas em alguns casos a ambigüidade continuou.

Problemas

PROBLEMAS BÁSICOS

1. a. Use o dicionário do código na Figura 9.6 para completar o quadro seguinte. Suponha que a leitura é da esquerda para a direita e que as colunas representam alinhamentos transcricionais e traducionais.

C										Dupla hélice de DNA
					T	G	A			
	C	A			U					mRNA transcrito
								G	C	A
										Anticódon apropriado de tRNA
				Trp						Aminoácidos incorporados à proteína

b. Marque as pontas 5' e 3' do DNA e RNA, bem como as pontas carboxila da proteína.

2. Considere os seguintes segmentos de DNA:

5'GCTTCCCAA 3'
3'CGAAGGGTT 5'

Suponha que o filamento de cima é o filamento-molde usado pela RNA polimerase.

a. Desenhe o RNA transcrito.

b. Marque suas pontas 5' e 3'.

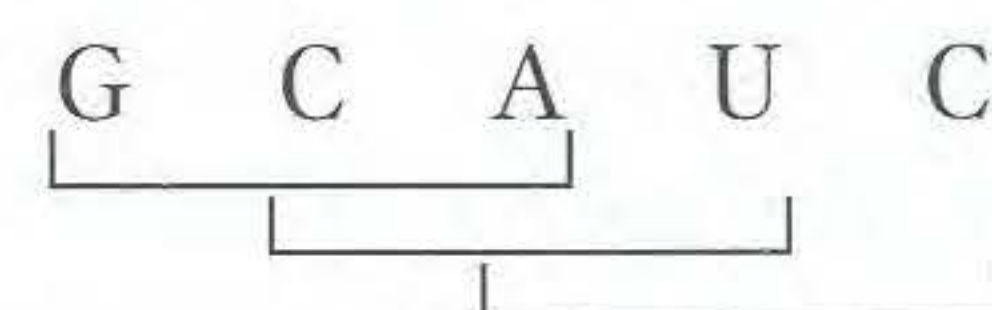
c. Desenhe a cadeia de aminoácidos correspondente.

d. Marque suas pontas amino e carboxila.

Repita as partes *a* até *d*, supondo que o filamento de baixo seja o filamento-molde.

3. Um evento mutacional insere um par de nucleotídeos extra no DNA. Qual dos seguintes resultados você espera? (1) Nenhuma proteína; (2) uma proteína na qual um aminoácido foi trocado; (3) uma proteína na qual três aminoácidos foram trocados; (4) uma proteína na qual dois aminoácidos foram trocados; (5) uma proteína na qual a maioria dos aminoácidos após o sítio de inserção foram mudados.

4. Antes que a verdadeira natureza do processo de código genético fosse totalmente compreendida, foi proposto que a mensagem podia ser lida em trincas superpostas. Por exemplo, a seqüência GCAUC pode ser lida como GCA CAU AUC:



Crie um teste experimental para essa idéia.

5. Nos sistemas de síntese de proteínas *in vitro*, a adição de um mRNA humano específico ao aparato traducional de *E. coli* (ribossomos, tRNA, e assim por diante) estimula a síntese de uma proteína muito parecida com a especificada pelo mRNA. O que mostra esse resultado?

6. Que anticódon você prevê para uma espécie de tRNA levando isoleucina? Há mais de uma resposta possível? Caso sim, cite as respostas alternativas.

7. **a.** Em quantos casos no código genético você não saberia o aminoácido especificado por um códon se você soubesse apenas os dois primeiros nucleotídeos do códon?

- b.** Em quantos casos você não sabe os dois primeiros nucleotídeos do códon se você soubesse qual o aminoácido especificado por ele?
- 8.** Deduza qual dos seis tipos selvagens de códons podem ter estado nos mutantes que levaram Brenner a deduzir a natureza do códon UAG.
- 9.** Se um polirribonucleotídeo contém quantidades iguais de bases adenina e uracil aleatoriamente posicionadas, que proporção de suas trincas codificará (a) fenilalanina, (b) isoleucina, (c) leucina, (d) tirosina?
- 10.** Você sintetizou três RNA mensageiros diferentes com bases incorporadas em sequência aleatória nas seguintes proporções: **(a)** 1 U:5 C, **(b)** 1 A:1 C:4 U, **(c)** 1 A:1 C:1 G:1 U. Em um sistema *in vitro* de síntese de proteínas, indique as identidades e as proporções dos aminoácidos que serão incorporadas nas proteínas quando cada um desses mRNA é testado. (Veja a Figura 9.6.)
- 11.** No fungo *Neurospora*, alguns mutantes foram obtidos que não tinham atividade de uma certa enzima. As mutações foram encontradas, por mapeamento, como sendo em um dentre dois genes não-ligados. Cite uma possível explicação em relação à estrutura quaternária da proteína.
- 12.** Foi encontrado um mutante que não tem todas as funções detectáveis para uma enzima específica. Se você tivesse um anticorpo marcado que detecta essa proteína em uma transferência Western (veja o Capítulo 1), você esperaria que houvesse alguma proteína detectável pelo anticorpo no mutante? Explique.
- 13.** Na transferência Western (veja o Capítulo 1), a enzima triptofano sintetase geralmente mostra duas bandas de mobilidade diferente no gel. Alguns mutantes sem atividade enzimática mostraram exatamente as mesmas bandas que o tipo selvagem. Outros mutantes sem atividade mostraram apenas a banda lenta; outros ainda, só a banda rápida.
- a.** Explique os tipos diferentes de mutantes a nível de estrutura de proteínas.
- b.** Por que você acha que não existem mutantes que não apresentem bandas?
- 14.** Nos experimentos de Crick-Brenner descritos neste capítulo, três “inserções” ou três “deleções” restauraram a matriz de leitura normal e a dedução foi que o código foi lido em grupos de três. Essa dedução foi realmente provada pelos experimentos? Pode um códon ser composto de seis bases, por exemplo?
- 15.** Um mutante não tem atividade para a enzima isocitrato liase. Esse resultado prova que a mutação é no gene codificante de isocitrato liase?
- 16.** Determinado supressor sem sentido corrige um mutante que não cresce para um estado que está próximo, mas não exatamente, ao tipo selvagem (ele tem crescimento anormal). Sugira um motivo possível para que a reversão não seja a correção total.
- 17.** Nos genes bacterianos, tão logo seja produzido um transcrito parcial de mRNA pelo sistema RNA polimerase, o ribossomo pula para ele e começa a tradução. Desenhe um diagrama desse processo, identificando as pontas 5' e 3' do mRNA, as pontas COOH e NH₂ da proteína, a RNA polimerase e pelo menos um ribossomo. (Por que esse sistema não funcionaria nos eucariontes?)
- 18.** Em um haplóide, um supressor sem sentido *su1* atua na mutação 1, mas não na mutação 2 ou 3 do gene *P*. Um supressor sem sentido não-ligado *su2* funciona na mutação 2 de *P*, mas não em 1 ou 3. Explique esse padrão de supressão com relação à natureza das mutações e supressores.
- 19.** Foram desenvolvidos sistemas de tradução *in vitro* nos quais moléculas específicas de RNA podem ser adicionadas a um tubo de ensaio contendo um extrato de células bacterianas que inclui todos os componentes necessários para a tradução (ribossomos, tRNA, aminoácidos). Se um aminoácido marcado radioativamente é incluído, qualquer proteína traduzida desse RNA pode ser detectada e apresentada em um gel. Se um mRNA eucariótico for adicionado ao tubo de ensaio, a proteína radioativa seria produzida? Explique.
- 20.** Em um sistema de tradução eucariótico (contendo um extrato de célula de uma célula eucariótica) comparável ao do Problema 19, uma proteína seria produzida por um RNA bacteriano? Caso não, por quê?
- 21.** Um sistema de tradução quimérico contendo a subunidade maior do ribossomo de *E. coli* e a subunidade menor de levedura (um eucarionte unicelular) seria capaz de funcionar na síntese de proteínas? Explique por que sim ou não.
- 22.** As mutações que trocam um único aminoácido no sítio ativo de uma enzima podem resultar na síntese de quantidades tipo selvagem de uma enzima inativa. Você pode pensar em outras regiões em uma proteína onde uma mudança em um único aminoácido poderia ter o mesmo resultado?
- 23.** Que evidência apóia a visão de que os RNA ribossômicos são um componente mais importante do ribossomo do que as proteínas ribossômicas?
- 24.** Explique por que os antibióticos, tais como eritromicina e Zithromax, que se ligam à subunidade maior do ribossomo não nos prejudicam.
- 25.** Por que os eucariontes multicelulares precisam ter centenas de genes codificantes de cinase?
- 26.** Nosso sistema imunológico faz muitas proteínas diferentes que nos protegem de infecções virais e bacterianas. As empresas de biotecnologia devem produzir grandes quantidades dessas proteínas imunes para teste humano e eventual venda ao público. Para esse fim, seus cientistas construíram culturas de células bacterianas ou humanas para expressar essas proteínas imunes. Explique por que essas proteínas isoladas de culturas bacterianas são geralmente inativas, enquanto as mesmas proteínas isoladas de culturas humanas são ativas (funcionais).

PROBLEMAS DESAFIADORES

27. A adição de um único nucleotídeo e a deleção de um único nucleotídeo distantes aproximadamente 15 sítios no DNA causam uma mudança na sequência da proteína de:

Lis-Ser-Pro-Ser-Leu-Asn-Ala-Ala-Lis

para

Lis-Val-His-His-Leu-Met-Ala-Ala-Lis

- a. Quais são as sequências de nucleotídeos antiga e nova do mRNA? (Use o dicionário do códon na Figura 9.6.)
b. Que nucleotídeo foi adicionado e qual foi deletado? (O Problema 27 é de W.D. Stansfield, *Theory and Problems of Genetics*. McGraw-Hill, 1969.)
28. Você está estudando um gene de *E. coli* que especifica uma proteína. Uma parte de sua sequência é

-Ala-Pro-Trp-Ser-Glu-Lis-Cis-His-

Você recupera uma série de mutantes para esse gene que não apresentam atividade enzimática. Isolando os produtos da enzima mutante, você encontra as seguintes sequências:

Mutante 1:

-Ala-Pro-Trp-Arg-Glu-Lis -Cis -His-

Mutante 2:

-Ala-Pro-

Mutante 3:

-Ala-Pro-Gli -Val -Lis -Asn-Cis -His-

Mutante 4:

-Ala-Pro-Trp-Fen-Fen-Tre -Ciz -His-

Qual a base molecular para cada mutação? Qual a sequência de DNA que especifica essa parte da proteína?

29. Hoje são conhecidos supressores de mudanças de matriz de leitura. Proponha um mecanismo para sua ação.
30. Considere o gene que especifica a estrutura da hemoglobina. Disponha os seguintes eventos na sequência mais provável na qual ocorreriam.
- a. É observada anemia.
b. A forma do sítio de ligação ao oxigênio é alterada.
c. Um códon incorreto é transcrito no mRNA de hemoglobina.
d. O ovócito (gameta feminino) recebe uma alta dose de radiação.
e. Um códon incorreto é gerado no DNA do gene de hemoglobina.
f. Uma mãe (técnica de raios X) acidentalmente fica em frente a um gerador de raios X.
g. Uma criança morre.
h. A capacidade de transporte de oxigênio do corpo é gravemente prejudicada.

i. O anticódon do tRNA que se alinha é de um tipo que traz um aminoácido não-adequado.

j. A substituição de um par de nucleotídeos ocorre no DNA do gene de hemoglobina.

31. Um mutante celular induzido é isolado de uma cultura de tecido de *hamster* devido à sua resistência à α -amanitina (veneno derivado de um fungo). A eletroforese mostra que o mutante tem uma RNA polimerase alterada; só uma banda eletroforética está em uma posição diferente da polimerase tipo selvagem. As células são supostamente diplóides. O que os resultados desse experimento lhe dizem sobre os modos de detectar mutantes recessivos em tais células?
32. Uma molécula bifilamentar de DNA com a sequência mostrada aqui produz, *in vivo*, um polipeptídeo que tem cinco aminoácidos de tamanho.

TAC ATG ATC ATT TCA CGG AAT TTC TAG CAT GTA
ATG TAC TAG TAA AGT GCC TTA AAG ATC GTA CAT

- a. Que filamento de DNA é transcrito e em que sentido?
b. Marque as pontas 5' e 3' de cada filamento.
c. Se ocorre uma inversão entre a segunda e terceira trincas das pontas da esquerda e direita, respectivamente, e o mesmo filamento de DNA é transcrito, que tamanho terá o polipeptídeo resultante?
d. Suponha que a molécula original está intacta e que o filamento de baixo é transcrito da esquerda para a direita. Cite a sequência de bases e marque as pontas 5' e 3' do anticódon que insere o quarto aminoácido no polipeptídeo nascente. Qual é esse aminoácido?
33. Uma das técnicas usadas para decifrar o código genético foi sintetizar polipeptídeos *in vitro*, com o uso de mRNA sintético com várias sequências de bases repetidas — por exemplo, (AGA)_n, que pode ser escrita como AGAAGAA-GAAGAAGA... Às vezes, o polipeptídeo sintetizado contém apenas um aminoácido (um homopolímero), e às vezes contém mais de um (um heteropolímero), dependendo da sequência repetida usada. Além disso, às vezes polipeptídeos diferentes foram feitos do mesmo mRNA sintético, sugerindo que o início da síntese de proteínas no sistema *in vitro* nem sempre começa em uma ponta do nucleotídeo do mensageiro. Por exemplo, podem ser feitos três polipeptídeos a partir de (AGA)_n: o homopolímero aa₁ (abreviado aa₁-aa₁), homopolímero aa₂ (aa₂-aa₂) e homopolímero aa₃ (aa₃-aa₃). Esses polipeptídeos provavelmente correspondem às seguintes leituras derivadas do começo em locais diferentes na sequência:

AGA AGA AGA AGA...
GAA GAA GAA GAA...
AAG AAG AAG AAG...

O quadro seguinte mostra os resultados reais obtidos do experimento feito por Khorana.

mRNA sintético	Polipeptídeo(s) sintetizado(s)
(UG) _n	(Ser-Leu)
(UG) _n	(Val-Cis)
(AC) _n	(Tre-His)
(AG) _n	(Arg-Glu)
(UUC) _n	(Ser-Ser) e (Leu-Leu) e (Fen-Fen)
(UUG) _n	(Leu-Leu) e (Val-Val) e (Cis-Cis)
(AAG) _n	(Arg-Arg) e (Lis-Lis) e (Glu-Glu)
(CAA) _n	(Tre-Tre) e (Asn-Asn) e (Gln-Gln)
(UAC) _n	(Tre-Tre) e (Leu-Leu) e (Tir-Tir)
(AUC) _n	(Ile-Ile) e (Ser-Ser) e (His-His)
(GUA) _n	(Ser-Ser) e (Val-Val)
(GAU) _n	(Asp-Asp) e (Met-Met)
(UAUC) _n	(Tir-Leu-Ser-Ile)
(UUAC) _n	(Leu-Leu-Tre-Tir)
(GAUA) _n	Nenhum
(GUAA) _n	Nenhum

Nota: A ordem na qual os polipeptídeos ou aminoácidos são listados no quadro não é significativa para (UAUC)_n e (UUAC)_n.

a. Por que (GUA)_n e (GAU)_n codificam apenas dois homopolipeptídeos?

b. Por que (GAUA)_n e (GUAA)_n falham em estimular a síntese?

c. Atribua um aminoácido a cada trinca na lista seguinte. Tenha em mente que, em geral, existem vários códons para um único aminoácido, e que a primeira das duas letras em um códon é geralmente a importante (mas a terceira letra é ocasionalmente significativa). Lembre também que alguns códons muito diferentes às vezes codificam o mesmo aminoácido. Tente fazer esta tarefa sem consultar a Figura 9.6.

AUG	GAU	UUG	AAC
GUG	UUC	UUA	CAA
GUU	CUC	AUC	AGA
GUA	CUU	UAU	GAG
UGU	CUA	UAC	GAA
CAC	UCU	ACU	UAG
ACA	AGU	AAG	UGA

A solução deste problema requer tanto lógica quanto ensaio-e-erro. Não desanime: Khorana recebeu um Prêmio Nobel para fazer isso. Boa sorte!
(O Problema 33 é de J. Kuspira and G.W. Walker, *Genetics: Questions and Problems*. McGraw-Hill, 1973.)

EXPLORANDO GENOMAS Um Tutorial de Bioinformática Baseado na Web

Encontrando Domínios Conservados

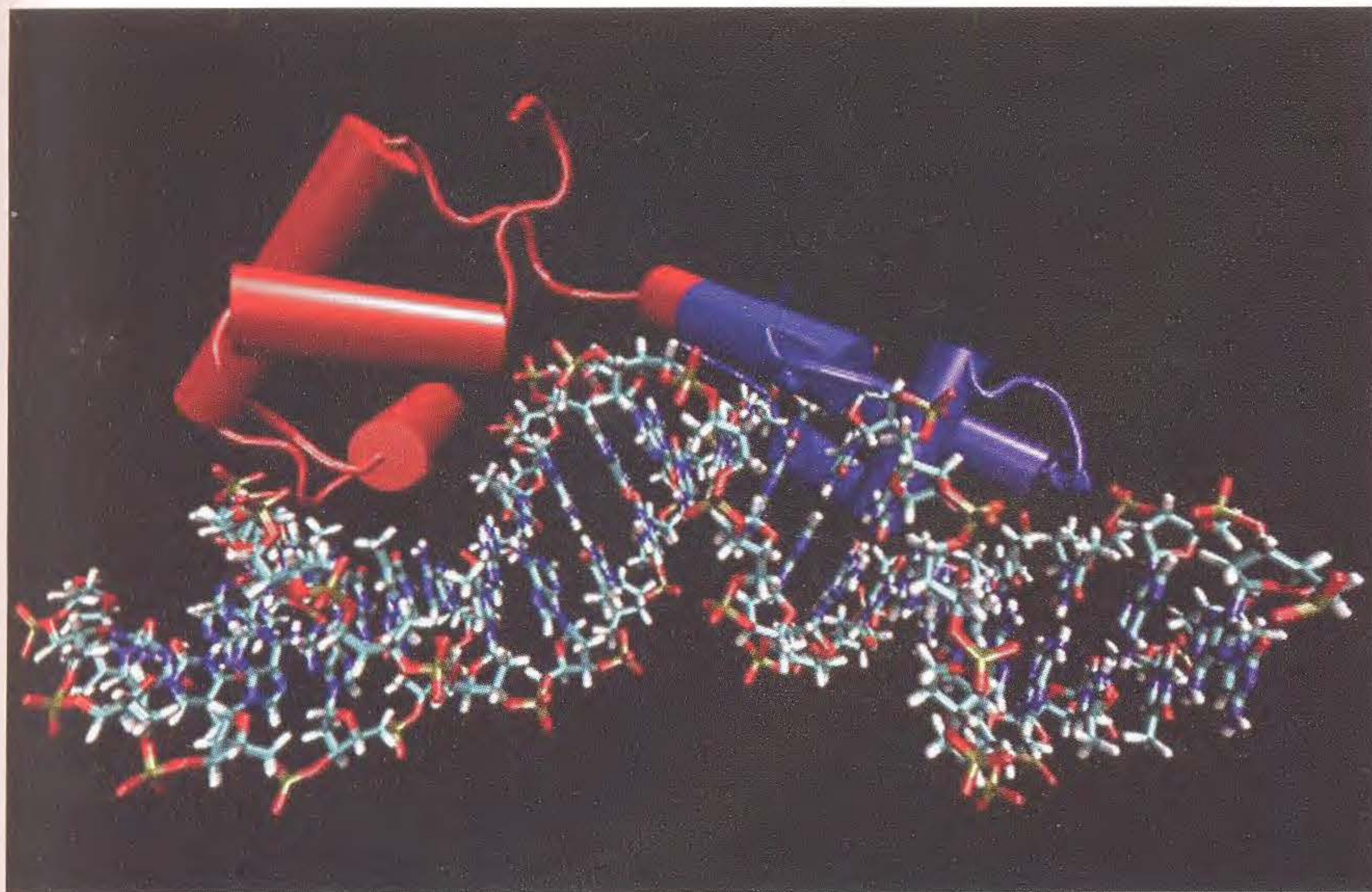
As seqüências de proteínas conservadas são manifestações da conservação de aminoácidos necessários para a estrutura, regulação ou função catalítica. Outros grupos de unidades podem ser identificados como sendo um padrão ou assinatura de um tipo particular de enzima ou domínio regulador. No tutorial de genômica em www.whfreeman.com/iga9e, você aprenderá como encontrar os domínios conservados em uma proteína complexa.

Determinando a Estrutura da Proteína

A função da proteína depende de uma estrutura tridimensional, a qual, por sua vez, depende da seqüência primária da proteína. A estrutura da proteína é determinada experimentalmente por cristalografia de raios X ou por ressonância magnética nuclear. Na ausência de informações experimentais diretas, são usados poderosos programas para tentar ajustar os dados da seqüência primária de aminoácidos a um modelo tridimensional. Tente um dos tutoriais de genômica em www.whfreeman.com/iga9e.

Regulação da Expressão Gênica em Bactérias e seus Vírus

10



O controle da expressão gênica é feito primariamente por proteínas de ligação ao DNA que reconhecem seqüências controladoras específicas dos genes. Aqui, a ligação da proteína repressora de Lac ao DNA do operador *lac* é modelada. [Cortesia do Dr. Timothy Paustian.]

Perguntas Fundamentais

- De que modos variam os níveis de transcrição gênica e tradução do RNA?
- Como as células ou vírus percebem mudanças ambientais e disparam mudanças na expressão gênica?
- Quais são os mecanismos moleculares de regulação gênica nas bactérias e seus vírus?
- Como é coordenada a expressão de conjuntos de genes?

Tópicos

- 10.1 Regulação gênica
- 10.2 Descoberta do sistema *lac*: controle negativo
- 10.3 Repressão catabólica do óperon *lac*: controle positivo
- 10.4 Controle duplo, positivo e negativo: o óperon de arabinose
- 10.5 Vias metabólicas e níveis adicionais de regulação: atenuação
- 10.6 Ciclos de vida de bacteriófagos: mais reguladores, óperons complexos
- 10.7 Fatores sigma alternativos regulam grandes conjuntos de genes

Em outubro de 1965, o rei da Suécia concedeu o Prêmio Nobel de Fisiologia ou Medicina a François Jacob, Jacques Monod e André Lwoff, do Instituto Pasteur, por suas descobertas de como a expressão gênica é regulada (Figura 10.1). Esse prêmio foi o fruto de uma excepcional colaboração de três homens extraordinários. Foi também um triunfo sobre grandes desafios. As chances de que esses três homens vivessem para ver esse dia, atingindo tal honraria, eram remotas.

Vinte e cinco anos antes, Monod, um estudante de doutorado na Sorbonne, em Paris, trabalhava sobre um fenômeno em bactérias, chamado de “adaptação enzimática”, que parecia tão obscuro para alguns que o diretor do laboratório de Zoologia, onde ele trabalhava, disse: “O que Jacques Monod está fazendo não tem nenhum interesse para a Sorbonne.” Jacob era um estudante de medicina de 20 anos de idade que desejava ser cirurgião. Lwoff, nessa época, era um membro bem estabelecido do Pasteur, chefe de seu departamento de Fisiologia Microbiana.

Veio então a Segunda Guerra Mundial.

Quando a França foi invadida e subjugada, Jacob correu para o litoral para se juntar às forças francesas que estavam reunindo-se na Inglaterra. Ele serviu como médico na África do Norte e na Normandia até ser muito ferido em um combate. Monod juntou-se à Resistência Francesa enquanto continuava seu trabalho. Após um ataque da Gestapo a seu laboratório na

Pioneiros da regulação gênica

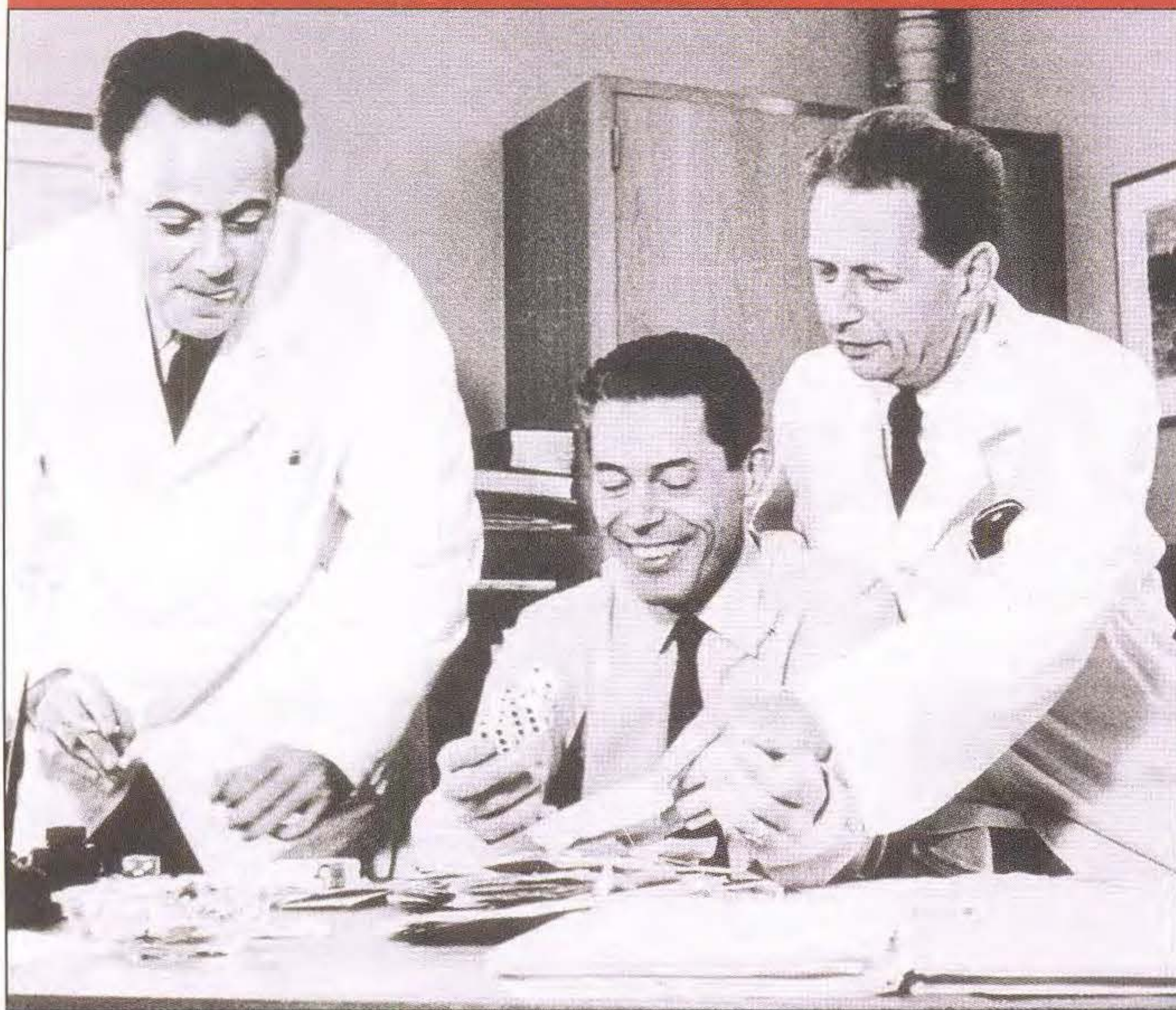


FIG. 10.1 François Jacob, Jacques Monod e André Lwoff foram premiados com o Prêmio Nobel de 1965 de Fisiologia ou Medicina por seu trabalho pioneiro sobre como é regulada a expressão gênica. [Instituto Pasteur.]

Sorbonne, ele concluiu que trabalhar ali era perigoso (seu predecessor na Resistência tinha sido preso e executado), e André Lwoff ofereceu-lhe um espaço no Pasteur. Monod, por sua vez, convenceu Lwoff a se juntar à Resistência, o que colocou o cientista em duplo risco: como judeu, Lwoff também se arriscou a ser deportado (15 cientistas do Pasteur foram deportados e, então, executados durante a ocupação da França).

Após a liberação de Paris, Monod serviu no exército francês e obteve, em uma biblioteca móvel do exército dos USA, um artigo de Oswald Avery *et al.* demonstrando que o DNA era o material hereditário em bactérias (veja o Capítulo 7). Seu interesse por genética reacendeu-se, e ele juntou-se a Lwoff após a guerra. Os ferimentos de Jacob foram muito graves para que ele fizesse carreira como cirurgião. Inspirado pelo enorme impacto dos antibióticos introduzidos mais tarde durante a guerra, Jacob decidiu fazer pesquisas científicas. Ele procurou Lwoff várias vezes pretendendo uma posição em seu laboratório, mas foi recusado. Fez então uma última tentativa e pegou Lwoff em um dia jovial. O cientista-chefe disse a Jacob: “Você sabe, nós acabamos de descobrir a indução do profago. Você teria interesse em trabalhar com o fago?” Jacob não tinha idéia do que Lwoff estava falando. Mas respondeu: “É exatamente o que eu gostaria de fazer.”

O grupo estava completo. O que foi revelado na década seguinte foi uma das mais criativas e produtivas colaborações na história da Genética, cujas descobertas ainda reverberam na Biologia de hoje.

Um dos mais importantes *insights* surgiu não no laboratório, mas em um cinema. Lutando com uma palestra

que ele tinha que preparar, Jacob optou por pegar sua esposa, Lise, e levá-la para uma matinê de domingo. Chateado e sonhando acordado, Jacob foi “envolvido por um súbito entusiasmo, misto de vago prazer... O espanto com o óbvio. Como não pensei nisto antes? Ambos os experimentos...no fago...e o feito com Pardee e Monod no sistema lactose...são idênticos! A mesma situação. O mesmo resultado...Em ambos os casos, um gene controla a formação de um produto citoplasmático, um repressor bloqueando a expressão de outros genes, e assim impedindo a síntese de galactosidase ou a multiplicação do vírus...Onde o repressor agiria para parar tudo subitamente? A única resposta simples...é no próprio DNA!” (F. Jacob, *The Statue Within: An Autobiography*, 1988).

Assim nasceu o conceito de um repressor agindo no DNA para reprimir a indução de genes. Levou muitos anos até que os hipotéticos repressores fossem isolados e caracterizados bioquimicamente. Os conceitos desenvolvidos por Jacob e Monod e explicados neste capítulo, o RNA mensageiro, promotores, operadores, genes reguladores, óperons e proteínas alostéricas, foram deduzidos inteiramente de evidências genéticas, e esses conceitos moldaram o futuro campo da Genética Molecular.

Walter Gilbert, que isolou o primeiro repressor e depois recebeu um Prêmio Nobel em Química por co-inventar um método de seqüenciamento do DNA, explicou o efeito do trabalho de Jacob e Monod naquela época: “A maioria das descobertas cruciais em ciência são de uma natureza tão simples que são muito difíceis de conceber sem passar pela experiência envolvida na descoberta... A sugestão de Jacob e Monod fez com que coisas muito obscuras ficassem muito simples” (H.F. Judson, *The Eighth Day of Creation: Makers of the Revolution in Biology*, 1979).

Os conceitos de Jacob e Monod iluminaram muito além das enzimas e vírus bacterianos. Eles compreenderam e foram capazes de articular, com excepcional eloquência, como suas descobertas sobre a regulação gênica estavam ligadas a mistérios gerais da diferenciação celular e desenvolvimento embrionário nos animais. Jacques Monod uma vez disse: “O que é verdade para *E. coli* também é verdade para o elefante.” Nos três próximos capítulos, veremos em que grau isso é verdade. Começaremos este capítulo com exemplos bacterianos que ilustram temas-chave e mecanismos na regulação da expressão gênica. Enfocaremos amplamente proteínas reguladoras simples e os “interruptores” genéticos nos quais elas atuam. Então, no Capítulo 11, abordaremos a regulação gênica e a diferenciação das células eucarióticas, que envolvem uma maquinaria bioquímica e genética mais complexa. Finalmente, no Capítulo 12, examinaremos o papel da regulação gênica no desenvolvimento de animais multicelulares. Veremos, então, como conjuntos de proteínas reguladoras atuam em grupos de interruptores genéticos para controlar a expressão gênica no tempo e no espaço, comandando a construção de corpos e partes do corpo.

10.1 Regulação Gênica

A despeito da simplicidade da forma, as bactérias têm em comum com membros maiores e mais complexos de outros reinos a tarefa fundamental de regular a expressão de seus genes. Um dos principais motivos é que elas são oportunistas nutricionais. Considere como as bactérias obtêm muitos compostos importantes, tais como açúcares, aminoácidos e nucleotídeos, necessários ao metabolismo. As bactérias nadam em um mar de potenciais nutrientes. Elas podem adquirir os compostos que precisam do ambiente, ou produzi-los pelas vias enzimáticas. A síntese das enzimas necessárias para essas vias gasta energia e recursos da célula; assim, dependendo da escolha, as bactérias obterão compostos do ambiente. Para serem econômicas, elas produzirão as enzimas necessárias para produzir esses compostos apenas quando não há outra opção, ou seja, quando esses compostos não estão disponíveis em seu ambiente local.

As bactérias desenvolveram sistemas reguladores que acoplam a expressão dos produtos gênicos a sistemas sensores que detectam o composto relevante em um ambiente bacteriano local. A regulação das enzimas que tomam parte no metabolismo de açúcar são um exemplo. As moléculas de açúcar podem ser oxidadas para produzir energia ou podem ser usadas como blocos estruturais para uma grande gama de compostos orgânicos. Entretanto, existem muitos tipos diferentes de açúcares que a bactéria pode usar, incluindo lactose, glicose, galactose e xilose. Uma importante proteína diferente é necessária para permitir que cada um dos açúcares entre na célula. Além disso, um conjunto diferente de enzimas é necessário para processar cada um dos açúcares. Se a célula fosse produzir simultaneamente todas as enzimas de que necessitasse, ela gastaria muita energia e materiais para produzir as enzimas de que poderia algum dia precisar para quebrar prospectivas fontes de carbono. A célula criou mecanismos que desligam (reprimem) a transcrição de todos os genes que codificam enzimas que não são necessárias em determinado tempo e ligar (ativar) esses genes que codificam enzimas agora necessárias. Por exemplo, se apenas lactose existe no ambiente, a célula irá desligar a transcrição de genes que codificam as enzimas necessárias para a importação e metabolismo de glicose, galactose, xilose e outros açúcares. Contrariamente, a célula iniciará a transcrição dos genes que codificam as enzimas necessárias para a importação e metabolismo de lactose. Em resumo, as células precisam de mecanismos que preenchem dois critérios:

1. Elas devem ser capazes de reconhecer condições ambientais nas quais elas devem ativar ou reprimir a transcrição de genes relevantes.
2. Elas devem ser capazes de ligar ou desligar, como um interruptor, a transcrição de cada gene específico ou grupo de genes.

Vamos criar o modelo atual para a regulação transcricional procariótica e, então, usar um exemplo bem compre-

endido — a regulação de genes no metabolismo do açúcar lactose — para examiná-lo em detalhe. Em particular, enfocaremos como esse sistema regulador foi dissecado com o uso das ferramentas da Genética Clássica e da Biologia Molecular.

As bases da regulação transcricional procariótica: interruptores genéticos

A regulação da transcrição depende principalmente de dois tipos de interações proteína-DNA. Ambas ocorrem perto do sítio no qual começa a transcrição gênica.

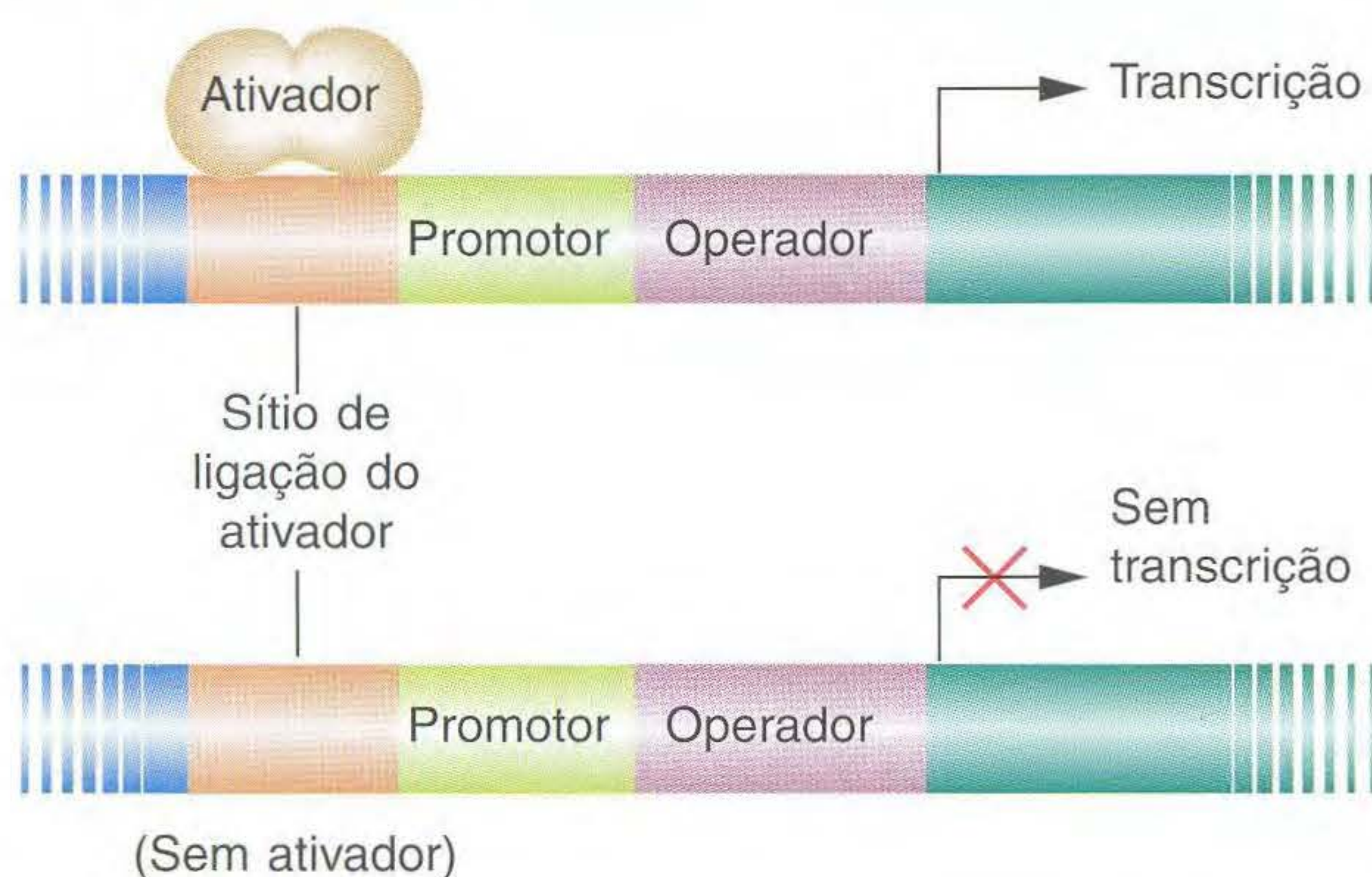
Uma dessas interações DNA-proteína determina onde começa a transcrição. O DNA que participa nessa interação é um segmento de DNA chamado de **promotor**, e a proteína que se liga a esse sítio é a RNA polimerase. Quando a RNA polimerase se liga ao DNA promotor, a transcrição pode começar algumas bases distantes do sítio promotor. Todos os genes devem ter um promotor ou não será transcrito.

O outro tipo de interação DNA-proteína decide se a transcrição ativada pelo promotor ocorre. Os segmentos de DNA perto do promotor servem como pontos de ligação para proteínas reguladoras específicas de seqüências chamadas **ativadores** e **repressores**. Nas bactérias, a maioria dos sítios de ligação para os repressores são chamados de **operadores**. Para alguns genes, uma proteína ativadora deve ligar-se a seu sítio-alvo no DNA como um pré-requisito para a transcrição começar. Tais casos às vezes são chamados de *regulação positiva* porque a *presença* da proteína ligada é necessária para a transcrição (Figura 10.2). Para outros genes, uma proteína repressora deve ser impedida de ligar-se a seu sítio-alvo como um pré-requisito para que a transcrição comece. Tais casos às vezes são chamados de *regulação negativa* porque a *ausência* do repressor ligado permite que a transcrição comece. Como os ativadores e repressores regulam a transcrição? Geralmente, uma proteína ativadora ligada ao DNA ajuda fisicamente a levar a RNA polimerase para seu promotor vizinho de modo que a polimerase possa começar a transcrição. Uma proteína repressora ligada ao DNA atua tipicamente seja interferindo fisicamente com a ligação da RNA polimerase a seu promotor (bloqueando o início da transcrição) ou impedindo o movimento da RNA polimerase ao longo da cadeia de DNA (bloqueando a transcrição). Juntas, essas proteínas regulatórias e seus sítios de ligação constituem **interruptores genéticos** que controlam as mudanças eficientes na expressão gênica dadas em resposta a condições ambientais.

Mudanças genéticas controlam a transcrição gênica. A função liga/desliga dos interruptores depende das interações de várias proteínas com seus sítios de ligação no DNA. A RNA polimerase interage com o promotor para começar a transcrição. As proteínas ativadoras ou repressoras ligam-se a sítios na vizinhança do promotor para controlar seu acesso à RNA polimerase.

As proteínas regulatórias controlam a transcrição

Regulação positiva



Regulação negativa

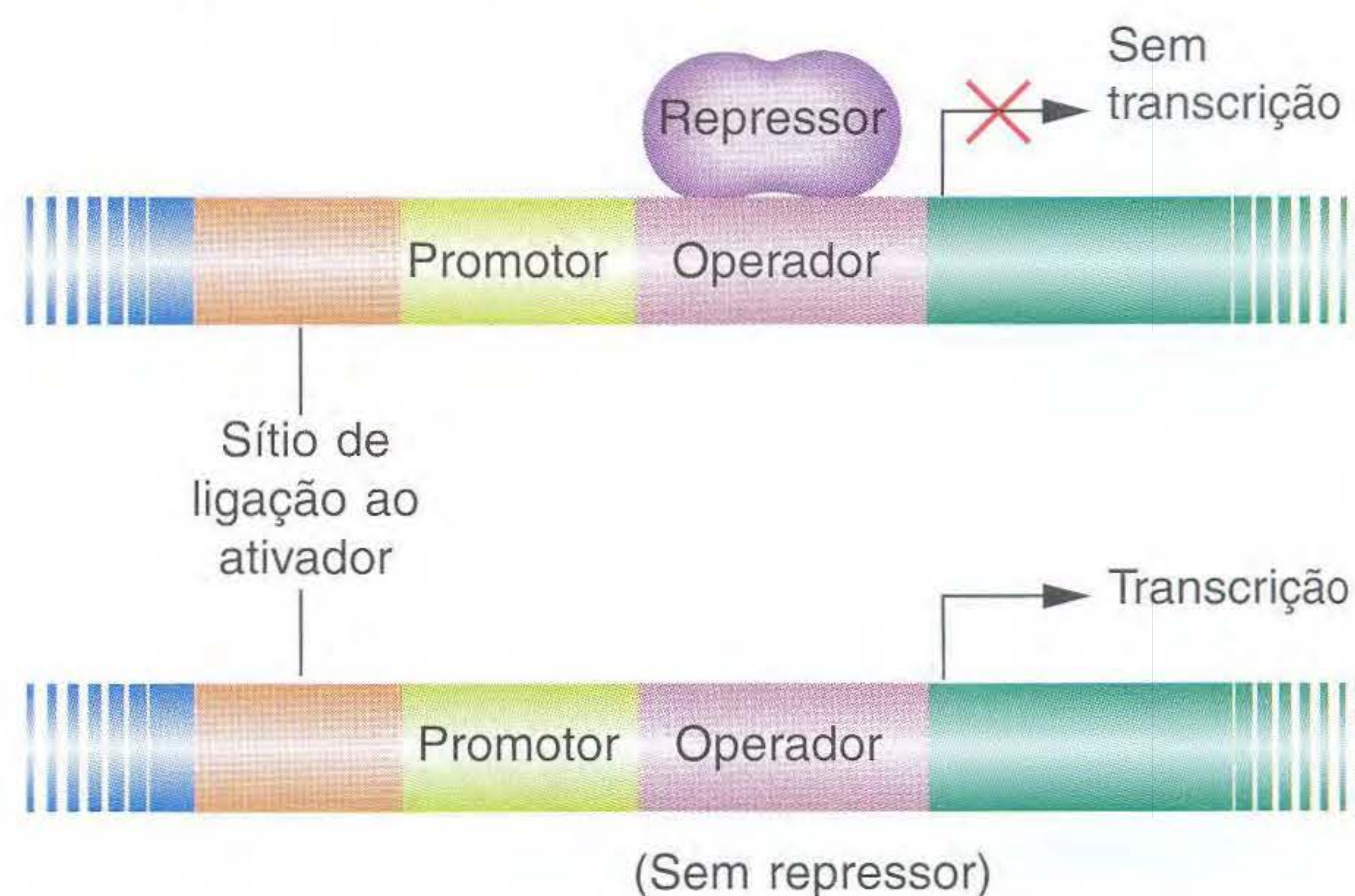


FIG. 10.2 A ligação de proteínas regulatórias pode ativar ou bloquear a transcrição.

Tanto as proteínas ativadoras quanto as repressoras devem ser capazes de reconhecer quando as condições ambientais são apropriadas para suas ações e agir de acordo. Assim, para que as proteínas ativadoras ou repressoras façam o seu trabalho, cada uma deve ser capaz de existir em dois estados: um que pode se ligar a seu DNA-alvo e outro que não pode. O estado de ligação deve ser apropriado para estabelecer as condições fisiológicas presentes e seu ambiente. Para muitas proteínas reguladoras, a ligação ao DNA é efetuada pela interação de dois sítios diferentes na estrutura tridimensional da proteína. Um sítio é o **domínio de ligação ao DNA**. O outro sítio, o **sítio alostérico**, atua como um sensor que ajusta o domínio de ligação ao DNA em um dos dois modos: funcional ou não-funcional. O sítio alostérico interage com pequenas moléculas chamadas *efetores alostéricos*. No metabolismo de lactose, um isômero do açúcar lactose (chamado de alolac-

tose) é um efector alostérico: o açúcar liga-se a uma proteína reguladora que inibe a expressão de genes necessários para o metabolismo de lactose. Em geral, um **efetor alostérico** liga-se ao sítio alostérico da proteína reguladora de tal modo que muda sua atividade. Nesse caso, a alolactose muda a estrutura do domínio de ligação ao DNA de uma proteína regulatória. Algumas proteínas ativadoras ou repressoras devem se ligar a seus efetores alostéricos antes que eles possam se ligar ao DNA. Outras só podem ligar-se ao DNA na ausência de seus efetores alostéricos. Duas dessas situações são mostradas na Figura 10.3.

Os efetores alostéricos controlam a habilidade das proteínas ativadoras ou repressoras em se ligar a seus DNA-alvo.

Os efetores alostéricos ligam-se a proteínas reguladoras

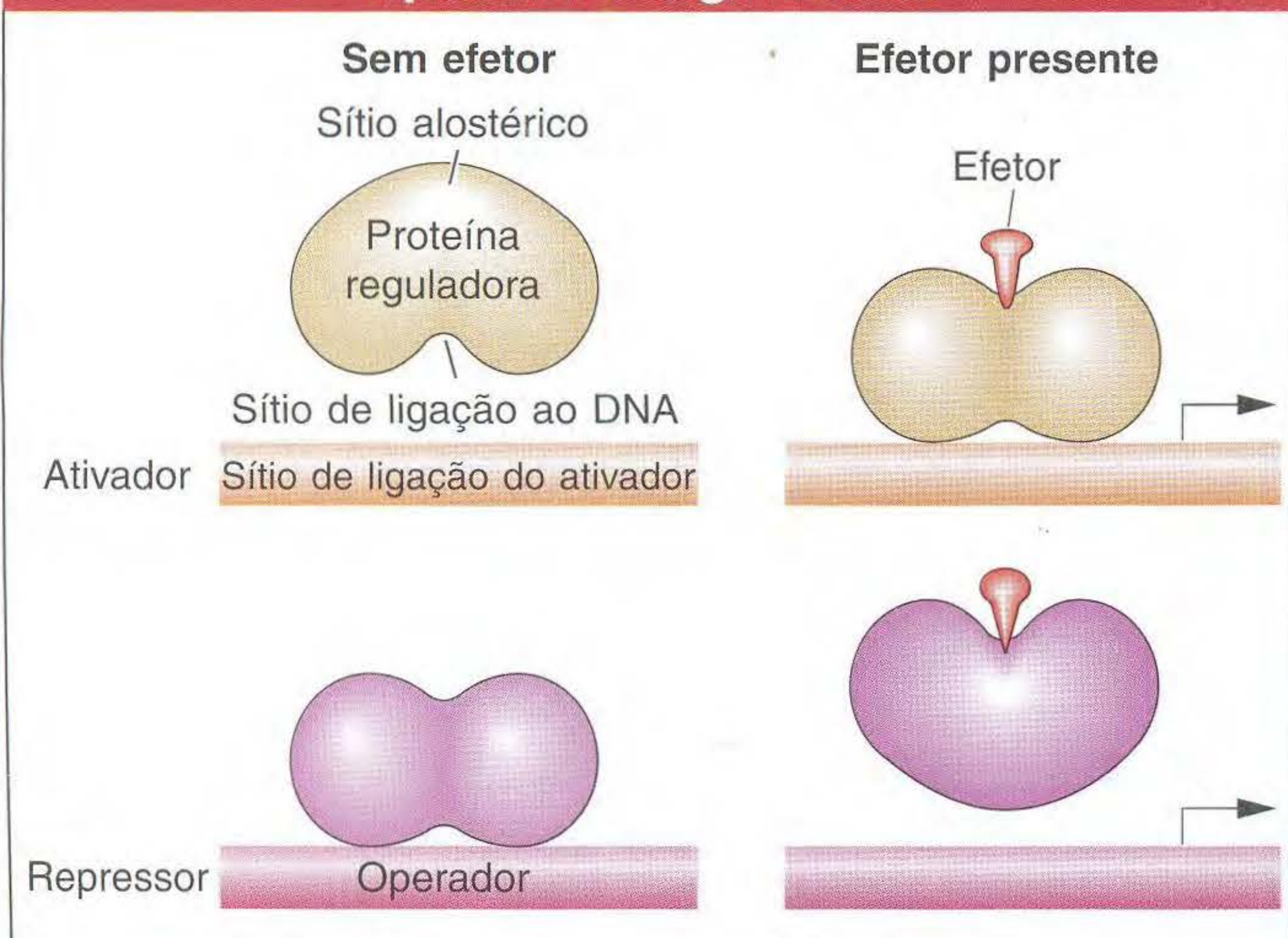


FIG. 10.3 Efetores alostéricos influenciam as atividades de ligação ao DNA de ativadores e repressores.

Um primeiro exame do circuito regulador *lac*

O trabalho pioneiro de François Jacob e Jacques Monod nos anos de 1950 mostrou como o metabolismo de lactose é regulado geneticamente. Vamos examinar o sistema em duas condições: a presença e a ausência de lactose. A Figura 10.4 é uma visão simplificada dos componentes desse sistema. O elenco de personagens da regulação do óperon *lac* inclui genes codificantes de proteínas e sítios no DNA que são alvos para as proteínas de ligação ao DNA.

Os genes estruturais *lac* O metabolismo de lactose requer duas enzimas: (1) uma permease, para transportar lactose para a célula, e (2) β -galactosidase, para quebrar a molécula de lactose para produzir glicose e galactose (Figura 10.5). As estruturas da β -galactosidase e permease são codificadas por duas seqüências adjacentes, Z e Y, respectivamente. Uma terceira seqüência contígua codifica uma enzima adicional, chamada *transacetilase*, que não é necessária para o metabolismo de lactose. Chamaremos os genes Z, Y e A de *genes estruturais* — em outras palavras, segmentos codificantes de proteínas, reservando o julgamento dessa categorização para

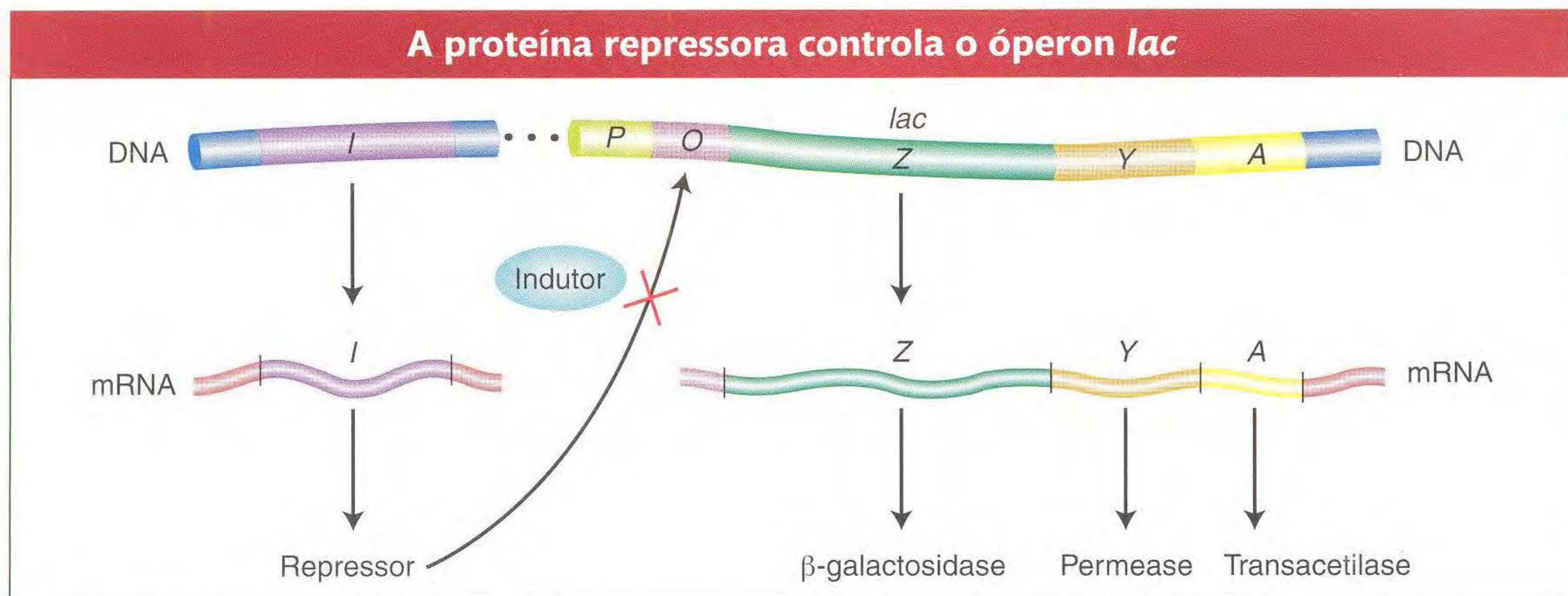


FIG. 10.4 Um modelo simplificado do óperon. A expressão coordenada dos genes Z, Y e A está sob controle negativo do produto do gene I, o repressor. Quando o indutor se liga ao repressor, o óperon é totalmente expresso.

mais adiante. Enfocaremos principalmente os genes Z e Y. Todos os três genes são transcritos em uma única molécula de RNA mensageiro. A regulação da produção desse mRNA coordena a síntese de todas as três enzimas. Isto é, todas ou nenhuma é sintetizada.

Se os genes que codificam proteínas constituem uma única unidade de transcrição, a expressão de todos esses genes será coordenadamente regulada.

Componentes regulatórios do sistema *lac* Os componentes regulatórios chave do sistema metabólico incluem um gene que codifica uma proteína regulatória da transcrição e dois sítios de ligação no DNA: um sítio para a proteína regulatória e outro sítio para a RNA polimerase.

1. *O gene para o repressor Lac.* Um quarto gene, o gene I, codifica a proteína repressora Lac, assim denominada porque pode bloquear a expressão dos genes Z, Y e A. O gene I está mapeado perto dos genes Z, Y e A, mas essa proximidade não parece ser importante para sua função.

2. *O sítio do promotor *lac*.* O promotor (P) é o sítio no DNA ao qual se liga a RNA polimerase para iniciar a transcrição dos genes estruturais *lac* (Z, Y e A).
3. *O sítio operador *lac*.* O operador (O) é o sítio do DNA ao qual se liga o repressor Lac. Ele está situado entre o promotor e o gene Z perto do ponto no qual começa a transcrição do mRNA multigênico.

A indução do sistema *lac* Os segmentos P, O, Z, Y e A (mostrados na Figura 10.6) constituem um **óperon**, definido como um segmento de DNA que codifica um mRNA multigênico, bem como um promotor comum adjacente e uma região regulatória. O genes *lacI*, codificando o repressor Lac, *não* é considerado parte do óperon *lac*, mas a interação entre o repressor Lac e o sítio operador *lac* é crucial para a regulação apropriada do óperon *lac*. O repressor Lac tem um sítio de *ligação ao DNA* que pode reconhecer a sequência do DNA operador e um *sítio alostérico* que liga lactose ou análogos de lactose que são úteis experimentalmente. O repressor irá ligar-se ao sítio do DNA perto dos genes que está controlando, e não a outros sítios distribuídos pelo cromossomo. Ligando-se ao operador, o repressor impede a transcrição pela RNA polimerase que se ligou ao sítio promotor adjacente; o óperon *lac* é “desligado”.

Quando a lactose ou seus análogos se ligam à proteína repressora, a proteína sofre uma **transição alostérica**, uma mudança de forma. Essa leve alteração na forma por sua vez altera o sítio de ligação ao DNA, de modo que o repressor não tem mais alta afinidade pelo operador. Assim, em resposta à ligação de lactose, o repressor sai do DNA: o óperon *lac* é “ligado”. A resposta do repressor à lactose satisfaz um requisito para tal sistema de controle, que a presença de lactose estimula a síntese dos genes necessários para seu processamento. O alívio da repressão por sistemas tais como o *lac* é chamado de **indução**. A lactose e seus análogos que inativam alostericamente o repressor e levam à expressão dos genes *lac* são chamados de **indutores**.

Vamos resumir como funciona o interruptor *lac*. Na ausência de um indutor (lactose ou um análogo), o repressor Lac liga-se ao sítio operador *lac* e impede a transcrição do ópe-

A lactose é degradada em dois açúcares

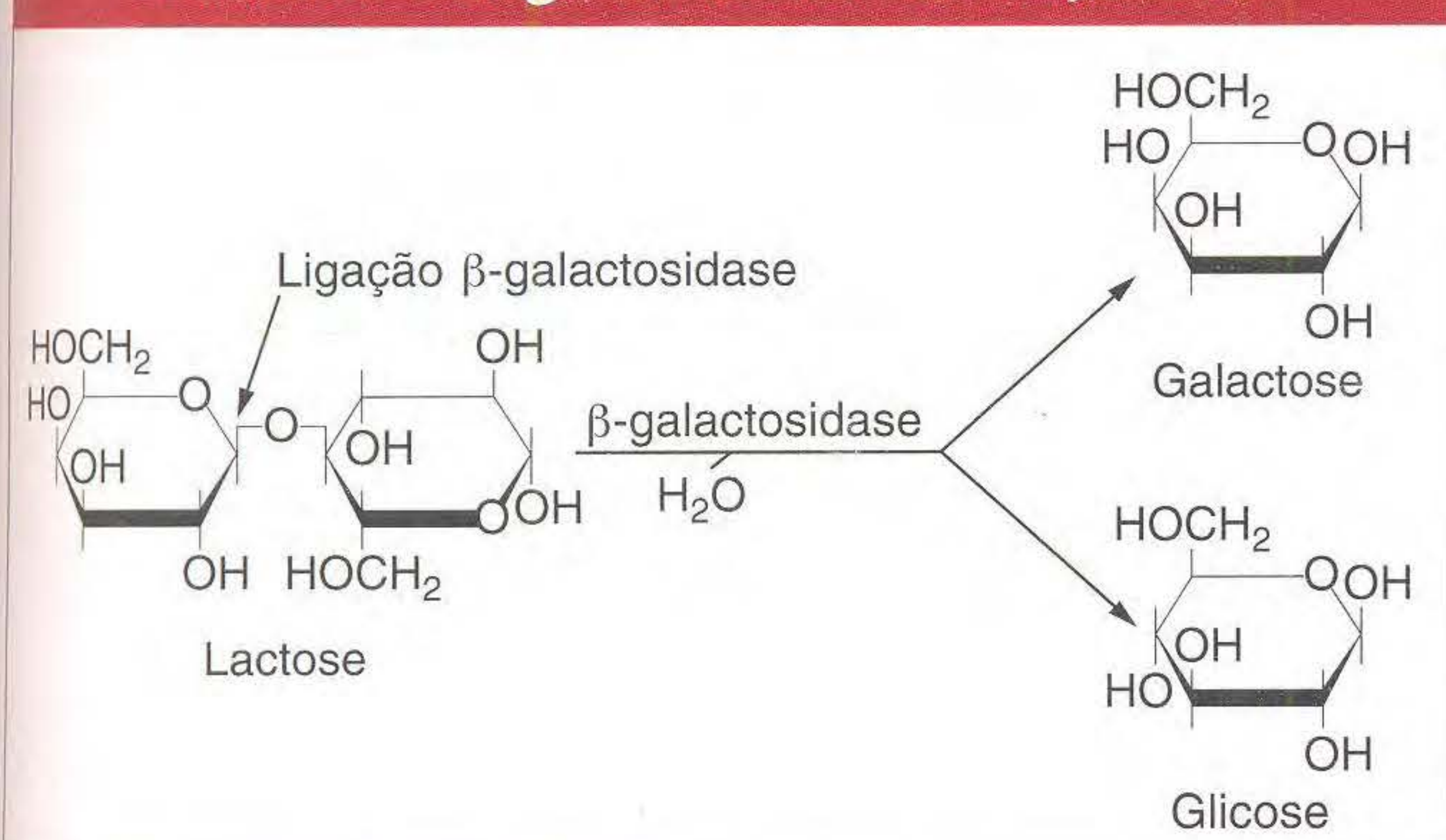
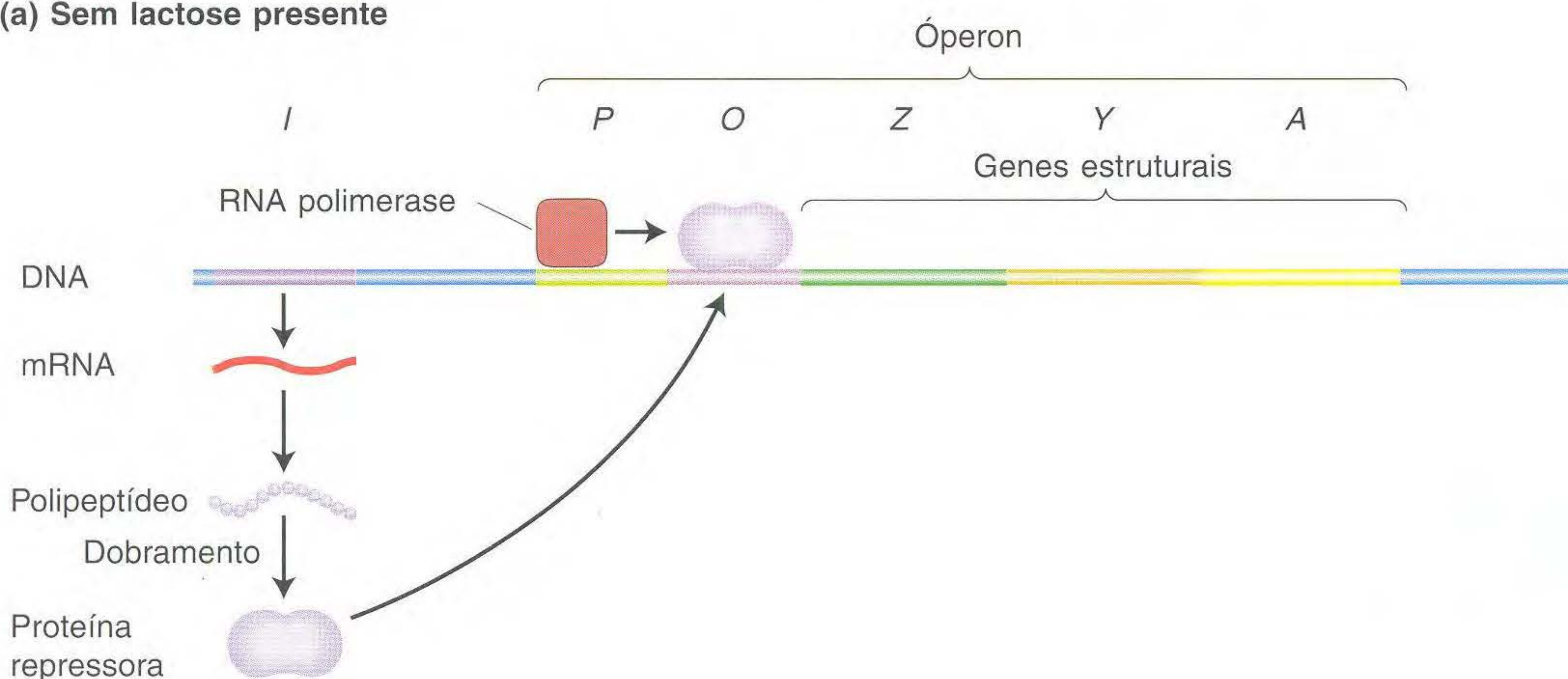


FIG. 10.5 O metabolismo da lactose. A enzima β-galactosidase catalisa uma reação na qual a água é adicionada à ligação β-galactosídica para quebrar a lactose em moléculas separadas de glicose e galactose.

O óperon *lac* é transcrito apenas na presença de lactose

(a) Sem lactose presente



(b) Lactose presente

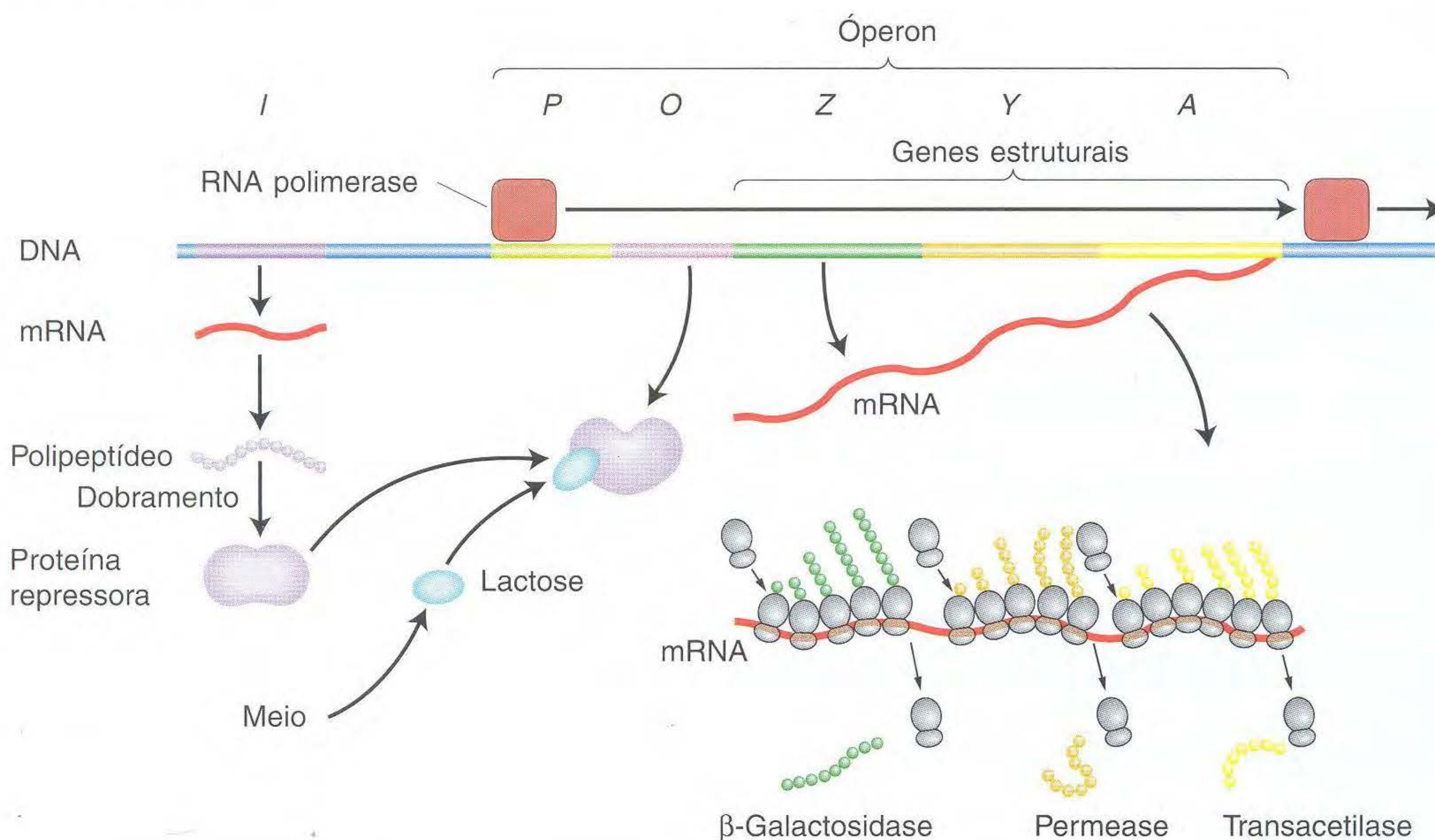


FIG. 10.6 Regulação do óperon *lac*. O gene *I* faz continuamente o repressor. (a) Na ausência de lactose, o repressor liga-se à região operadora *O* (operador) e bloqueia a transcrição. (b) A ligação de lactose muda a forma do repressor; logo, o repressor não se liga mais a *O* e sai do DNA. A RNA polimerase é então capaz de transcrever os genes estruturais *Z*, *Y* e *A*, e, assim, as três enzimas são produzidas.

ron *lac* bloqueando o movimento da RNA polimerase. Nesse sentido, o repressor Lac atua como um bloqueio no DNA. Conseqüentemente, todos os genes estruturais do óperon *lac* (os genes *Z*, *Y* e *A*) são reprimidos, e não há β-galactosidase, β-galactosídeo permease ou transacetilase na célula. Em contraste, quando um indutor está presente, ele se liga ao sítio alostérico de cada subunidade repressora Lac, inativando assim o sítio que se liga ao operador. O repressor Lac sai do DNA, permitindo que comece a transcrição dos genes estruturais do óperon *lac*. As enzimas β-galactosidase,

β-galactosídeo permease e transacetilase agora aparecem na célula de um modo coordenado.

► 10.2 Descoberta do Sistema *lac*: Controle Negativo

Para estudar a regulação gênica, idealmente precisamos de três coisas: uma análise bioquímica que nos deixe medir a quantidade de mRNA ou proteína expressa, ou ambos; de condições confiáveis nas quais os níveis de expressão dife-

rem em um genótipo tipo selvagem; e de mutações que perturbem os níveis de expressão. Em outras palavras, precisamos de um modo para descrever a regulação gênica tipo selvagem e precisamos de mutações que possam perturbar o processo regulatório tipo selvagem. Com esses elementos em mão, podemos analisar a expressão nos genótipos mutantes, tratando as mutações isoladamente e em combinação, para revelar qualquer tipo de evento de regulação gênica. A aplicação clássica desse enfoque foi usada por Jacob e Monod, que fizeram os estudos definitivos da regulação gênica bacteriana.

Jacob e Monod usaram o sistema de metabolismo de lactose de *E. coli* (veja Figura 10.4) para dissecar geneticamente o processo de indução enzimática — isto é, o aparecimento de uma enzima específica apenas na presença de seus substratos. Esse fenômeno foi observado nas bactérias por muitos anos, mas como uma célula podia “saber” exatamente que enzimas produzir? Como determinado substrato podia induzir o surgimento de uma enzima específica?

No sistema *lac*, a presença do indutor lactose faz com que as células produzam mais de 1.000 vezes a enzima β -galactosidase do que produziam quando cultivadas na ausência de lactose. Que papel o indutor tem no fenômeno da indução? Uma idéia era que o indutor fosse simplesmente ativador de um precursor de β -galactosidase que tinha se acumulado na célula. Entretanto, quando Jacob e Monod seguiram o destino de aminoácidos marcados radioativamente adicionados a células em crescimento antes ou depois da adição de um indutor, eles descobriram que a indução resultava na síntese de novas moléculas de enzimas, como indicado pela presença de aminoácidos radioativos nas enzimas. Essas novas moléculas podiam ser detectadas tão cedo quanto 3 min após a adição de um indutor. Além disso, a ausência do indutor causava uma parada abrupta na síntese de uma enzima. Portanto, ficou claro que a célula tem um mecanismo rápido e efetivo para ligar e desligar a expressão de um gene em resposta a sinais ambientais.

Genes controlados juntos

Quando Jacob e Monod induziram a β -galactosidase, eles descobriram que também induziram a enzima permease, que é necessária para o transporte de lactose para a célula. A análise de mutantes indica que cada enzima era codificada por um gene diferente. A enzima transacetilase (com uma função dispensável e ainda não conhecida) também era induzida junto com β -galactosidase e permease, e também mostrada depois sendo codificada por um gene separado. Portanto, Jacob e Monod puderam identificar três **genes controlados coordenadamente**. O mapeamento de recombinação mostrou que os genes *Z*, *Y* e *A* eram proximamente ligados no cromossomo.

Evidência genética do operador e repressor

Agora chegamos ao cerne do trabalho de Jacob e Monod: Como eles deduziram os mecanismos de regulação gênica no sistema *lac*? Seu enfoque foi o da Genética Clássica: examinar as consequências fisiológicas das mutações. Como veremos,

as propriedades das mutações nos genes estruturais e os elementos regulatórios do operon *lac* são bem diferentes, dando indícios importantes para Jacob e Monod.

Os indutores naturais, tais como a lactose, não são ótimos para esses experimentos, porque são hidrolisados pela β -galactosidase; a concentração do indutor diminui durante o experimento, e, assim, as dosagens da indução enzimática tornam-se bem complicadas. Em vez disso, para tais experimentos, Jacob e Monod usaram indutores sintéticos, tais como isopropil- β -D-tiogalactosídeo (IPTG; Figura 10.7). IPTG não é hidrolisado pela β -galactosidase.

Jacob e Monod descobriram que várias classes diferentes de mutações podem alterar a expressão dos genes estruturais do operon *lac*. Eles estavam interessados em avaliar as interações de novos alelos e, assim, saber que alelos exibiam dominância. Mas, para fazer tais testes, são necessários diplóides, e as bactérias são haplóides. Entretanto, Jacob e Monod foram capazes de produzir bactérias que são parcialmente diplóides inserindo fatores F' (veja o Capítulo 5) levando a região *lac* do genoma. Eles puderam então criar linhagens que eram heterozigotas para mutações *lac* selecionadas. Esses **diplóides parciais** permitiram que Jacob e Monod distinguíssem mutações no sítio regulatório de DNA (o operador *lac*) de mutações na proteína reguladora (o repressor Lac codificado pelo gene *I*).

Começamos examinando mutações que inativam genes estruturais para β -galactosidase e permease (designados Z^- e Y^- , respectivamente). A primeira coisa que aprendemos é que Z^- e Y^- são recessivos em relação a seus alelos tipo selvagem (Z^+ e Y^+). Por exemplo, a linhagem 2 no Quadro 10.1 pode ser induzida a sintetizar β -galactosidase (como a linhagem haplóide tipo selvagem 1 nesse quadro), muito embora seja heterozigota para os alelos *Z* mutantes e tipo selvagem. Isso demonstra que o alelo Z^+ é dominante em relação à sua contraparte Z^- .

Jacob e Monod primeiro identificaram duas classes de mutações reguladoras, chamadas O^C e I^- , que foram chamadas de **mutações constitutivas**, significando que elas fizeram com que os genes estruturais do operon *lac* fossem expressos independente de o indutor estar presente. Jacob e Monod identificaram a existência do operador com base na análise das mutações O^C . Essas mutações tornam o operador incapaz de se ligar ao repressor; elas danificam o interruptor de modo que ele fique sempre “ligado” (Quadro 10.1,

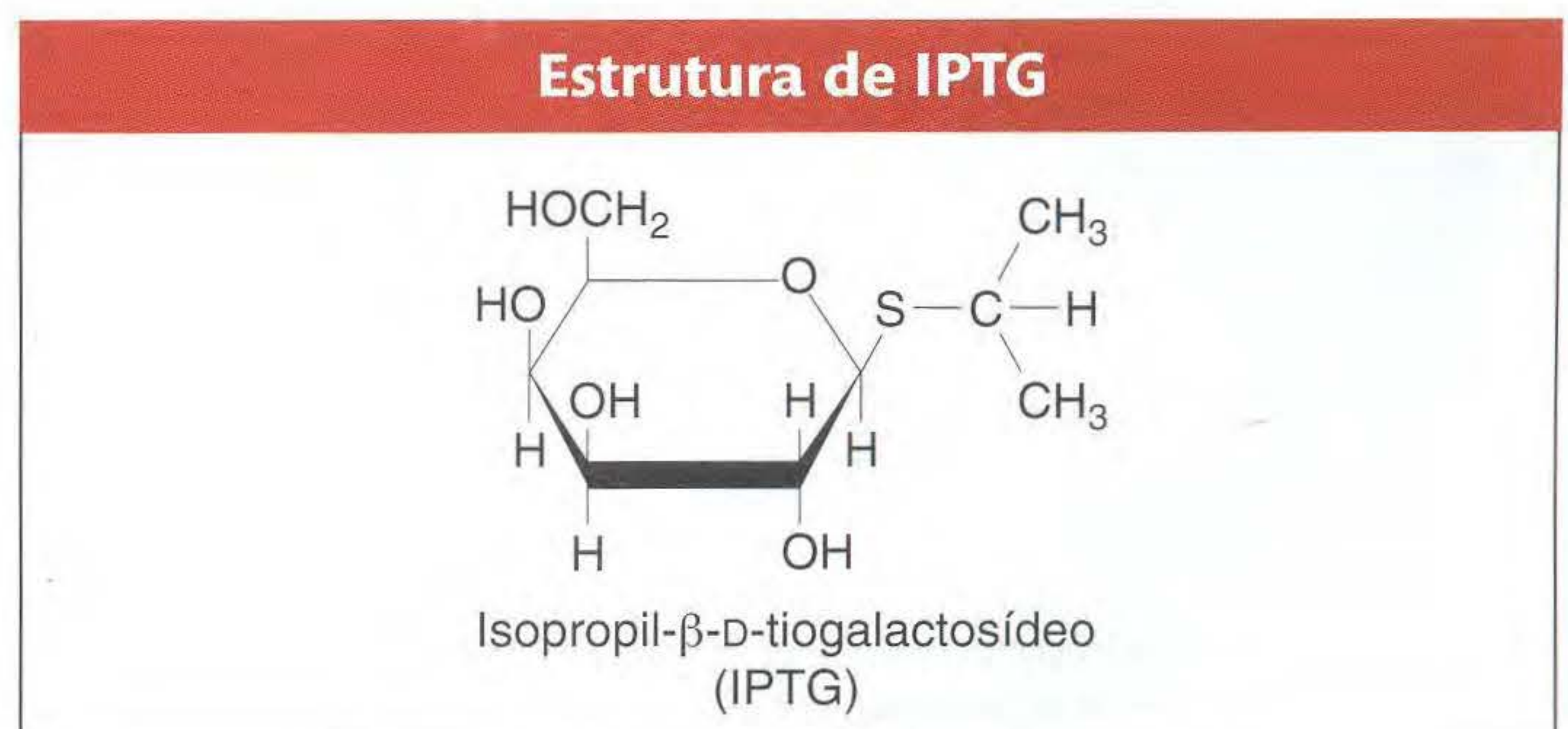


FIG. 10.7 IPTG é um indutor do operon *lac*.

Quadro 10.1 Síntese de β-galactosidase e Permease em Haplóides e Mutantes Operadores Diplóides Heterozigotos

Linhagem	Genótipo	β-galactosidase (Z)		Permease (Y)		Conclusão
		Não-induzida	Induzida	Não-induzida	Induzida	
1	O ⁺ Z ⁺ Y ⁺	-	+	-	+	Tipo selvagem é indutível
2	O ⁺ Z ⁺ Y ⁺ /F' ['] O ⁺ Z ⁻ Y ⁺	-	+	-	+	Z ⁺ é dominante em relação a Z ⁻
3	O ^c Z ⁺ Y ⁺	+	+	+	+	O ^c é constitutivo
4	O ⁺ Z ⁻ Y ⁺ /F' ['] O ^c Z ⁺ Y ⁻	+	+	-	+	Operador é de ação cis

Nota: As bactérias foram cultivadas em glicerol (sem glicose presente), com e sem o indutor IPTG. A presença ou a ausência da enzima são indicadas por + ou -, respectivamente. Todas as linhagens são I⁺.

linhagem 3). Importante, os efeitos constitutivos das mutações O^c foram restritos apenas aos genes estruturais *lac* no mesmo cromossomo que a mutação O^c. Por esse motivo, o operador mutante foi dito de **ação cis**, como demonstrado pelo fenótipo da linhagem 4 no Quadro 10.1. Aqui, como o gene de permease tipo selvagem (Y⁺) está em cis em relação ao operador tipo selvagem, a atividade da permease é induzida apenas quando a lactose ou um análogo está presente. Em contraste, o gene de β-galactosidase tipo selvagem (Z⁺) é cis em relação ao operador mutante O^c; logo, a β-galactosidase é expressa constitutivamente. Essa propriedade incomum de ação cis sugeriu que o operador é um segmento de DNA que influencia apenas a expressão dos genes estruturais ligados a ele (Figura 10.8). O operador, portanto, age simplesmente como um sítio de ligação à proteína e não faz um produto gênico.

Jacob e Monod fizeram testes genéticos comparáveis com as mutações I⁻ (Quadro 10.2). Uma comparação do tipo selvagem indutível I⁺ (linhagem 1) com as linhagens I⁻ mostra que as mutações I⁻ são constitutivas (linhagem 2). A linhagem 3 demonstra que o fenótipo indutível de I⁺ é dominante em

relação ao fenótipo constitutivo de I⁻. Esse achado mostrou a Jacob e Monod que a quantidade de proteína tipo selvagem codificada por uma cópia do gene é suficiente para regular ambas as cópias do operador em uma célula diplóide. Mais significativa, a linhagem 4 mostrou a eles que o produto do gene I⁺ é de **ação trans**, significando que o produto gênico pode regular todos os genes estruturais do óperon *lac*, estejam eles residindo na mesma molécula de DNA ou em diferentes (em cis ou trans, respectivamente). Ao contrário do operador, a ação do gene I é de um gene padrão codificante de proteínas. O produto proteico do gene I é capaz de se difundir através de uma célula e agir em ambos os operadores no diplóide parcial (Figura 10.9).

As mutações no operador revelam que tal sítio é de ação cis; isto é, regula a expressão de uma unidade adjacente de transcrição na mesma molécula de DNA. Em contraste, as mutações no gene codificante de uma proteína repressora revelam que essa proteína é de ação trans, isto é, podem atuar em qualquer cópia do sítio-alvo do DNA na célula.

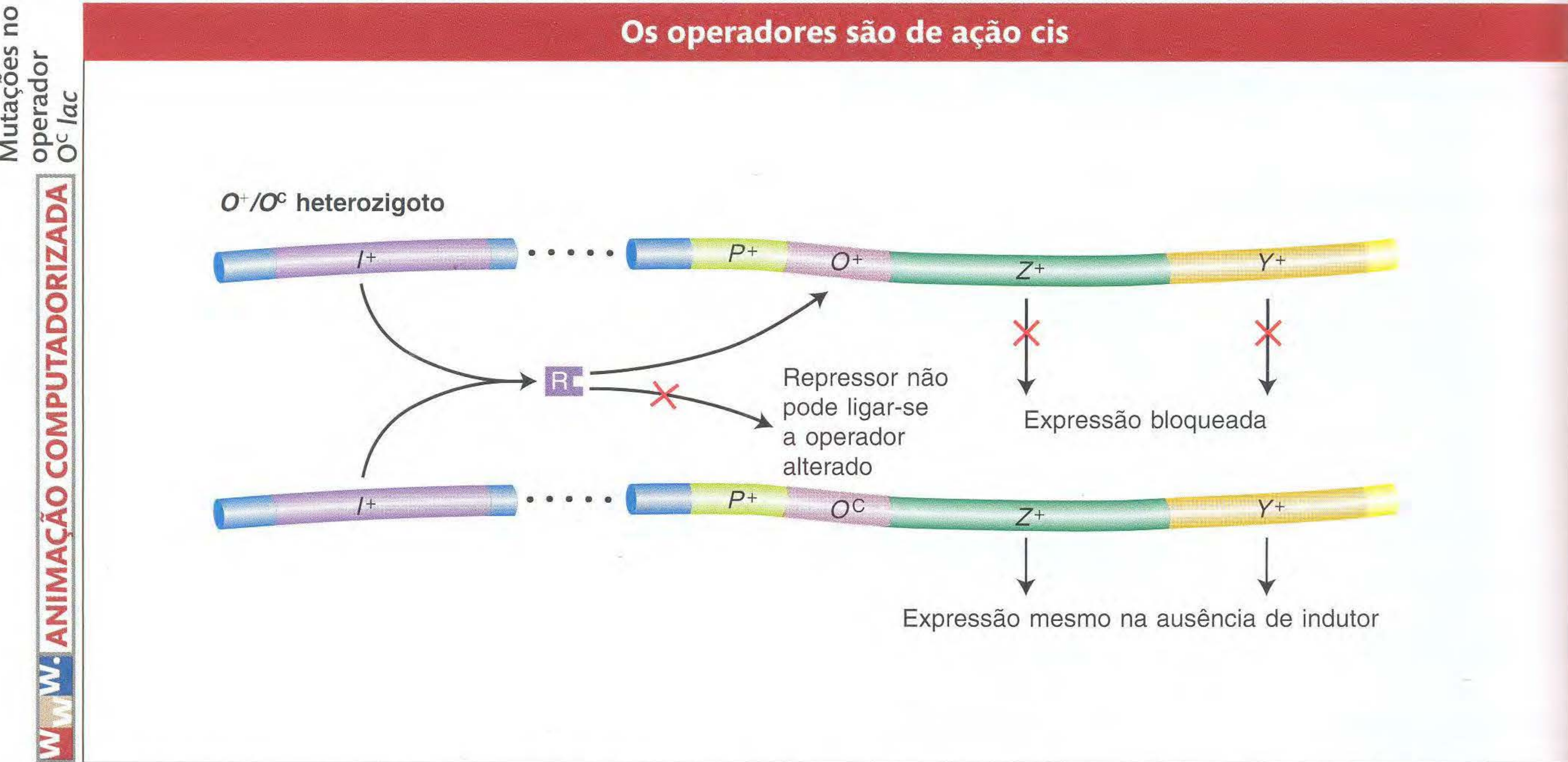


FIG. 10.8 Os heterozigotos O⁺/O^c demonstram que os operadores são de ação cis. Como um repressor não pode ligar-se a operadores O^c, os genes estruturais *lac* ligados a um operador O^c são expressos mesmo na ausência de um indutor. Entretanto, os genes *lac* adjacentes a um operador O⁺ ainda estão sujeitos a repressão.

Quadro 10.2 Síntese de β-galactosidase e Permease em Linhagens Haplóides e Diplóides Levando I⁺ e I⁻

Linhagem	Genótipo	β-galactosidase (Z)		Permease (Y)		Conclusão
		Não-induzida	Induzida	Não-induzida	Induzida	
1	I ⁺ Z ⁺ Y ⁺	-	+	-	+	I ⁺ é indutível
2	I ⁻ Z ⁺ Y ⁺	+	+	+	+	I ⁻ é constitutivo
3	I ⁺ Z ⁻ Y ⁺ /F ['] I ⁻ Z ⁺ Y ⁺	-	+	-	+	I ⁺ é dominante em relação a I ⁻
4	I ⁻ Z ⁻ Y ⁺ /F ['] I ⁺ Z ⁺ Y ⁻	-	+	-	+	I ⁺ é de ação trans

Nota: As bactérias foram cultivadas em glicerol (sem glicose presente) e induzidas com IPTG. A presença do nível máximo da enzima é indicada por um sinal mais; a ausência ou nível muito baixo de uma enzima é indicado por um sinal menos. (Todas as linhagens são O⁺.)

Evidência genética da alosteria

Finalmente, Jacob e Monod foram capazes de demonstrar alosteria pela análise de outra classe de mutações repressoras. Lembre que o repressor Lac inibe a transcrição do operon lac na ausência de um indutor, mas permite a transcrição quando o indutor está presente. Essa regulação é feita por um segundo sítio na proteína repressora, o sítio alostérico, que se liga ao indutor. Quando ligado ao indutor, o repressor sofre uma mudança na estrutura geral de tal modo que seu sítio de ligação ao DNA não pode mais funcionar.

Jacob e Monod isolaram uma outra classe de mutação repressora, chamada de mutações super-repressoras (I^s). As mutações I^s causam repressão mesmo na presença de um indutor (compare a linhagem 2 no Quadro 10.3 com a linhagem tipo selvagem I indutível). Ao contrário das mutações I⁻, as mutações I^s são dominantes em relação a I⁺ (veja Quadro 10.3, linhagem 3). Essa observação importante levou Jacob e Monod a especular que as mutações I^s alteram o sítio alostérico de modo que ele não pode mais se ligar a um indutor. Como uma consequência, a proteína repressora codificada por I^s liga-se continuamente ao operador,

evitando a transcrição do operon lac mesmo quando o indutor está presente na célula. Com base nisso, podemos ver por que I^s é dominante em relação a I⁺. A proteína mutante I^s irá ligar-se a ambos os operadores na célula, mesmo na presença de um indutor e independentemente de a proteína codificada por I⁺ poder estar presente na mesma célula (Figura 10.10).

Análise genética do promotor lac

A análise mutacional também demonstrou que um elemento essencial para a transcrição de lac está situado entre I e O. Esse elemento, chamado de promotor (P), serve como o sítio de iniciação para a transcrição, como descrito no Capítulo 7. Existem duas regiões de ligação para a RNA polimerase em um típico promotor procariótico, mostrado na Figura 10.11 como as duas regiões altamente conservadas em -35 e -10. As mutações promotoras são de ação cis, pois afetam a transcrição de todos os genes estruturais adjacentes no operon. Como os operadores e outros elementos de ação cis, os promotores são sítios na molécula de DNA que são ligados a proteínas, e eles mesmos não produzem produto proteico.

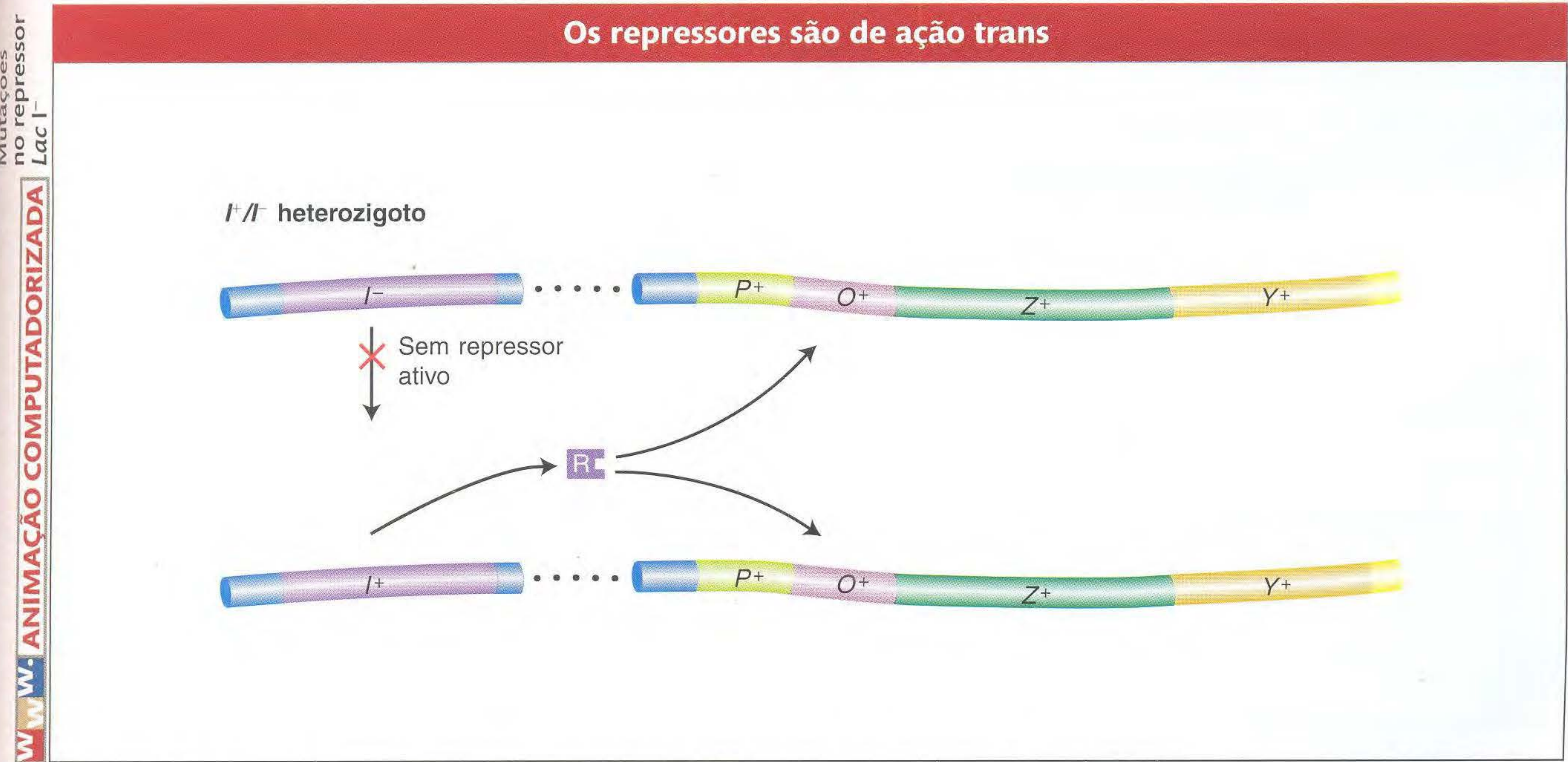


FIG. 10.9 A natureza recessiva das mutações I⁻ demonstra que o repressor é de ação trans. Embora nenhum repressor ativo seja sintetizado pelo gene I⁻, o gene tipo selvagem (I⁺) fornece um repressor funcional que se liga a ambos os operadores em uma célula diplóide e bloqueia a expressão do operon lac (na ausência de um indutor).

QUADRO 10.3 Síntese de β-galactosidase e Permease pelos Tipos Selvagens e por Linhagens Levando Alelos Diferentes do Gene *I*

Linhagem	Genótipo	β-galactosidase (Z)		Permease (Y)		Conclusão
		Não-induzida	Induzida	Não-induzida	Induzida	
1	$I^+Z^+Y^+$	-	+	-	+	I^+ é indutível
2	$I^sZ^+Y^+$	-	-	-	-	I^s é sempre reprimida
3	$I^sZ^+Y^+/F'I^+Z^+Y^+$	-	-	-	-	I^s é dominante em relação a I^+

Nota: As bactérias foram cultivadas em glicerol (sem glicose presente) com e sem o indutor IPTG. Presença da enzima indicada é representada por +, ausência ou níveis baixos, por -.

Caracterização molecular do repressor Lac e operador *lac*

Walter Gilbert e Benno Müller-Hill deram uma demonstração decisiva do sistema *lac*, em 1966, monitorando a ligação do indutor marcado radioativo IPTG a uma proteína repressora purificada. Eles primeiro mostraram que o repressor consiste em quatro subunidades idênticas e, portanto, contém quatro sítios de ligação de IPTG. Segundo, eles mostraram que, no tubo de ensaio, a proteína repressora liga-se ao DNA contendo o operador, e saem do DNA na presença de IPTG. (Uma descrição mais detalhada de como o repressor e outras proteínas de ligação ao DNA funcionam é dada depois, no final da Seção 10.6.)

Gilbert *et al.* mostraram que o repressor pode proteger bases específicas no operador dos reagentes químicos. Essa informação lhes permitiu isolar e determinar a sequência do DNA que constitui o operador. Eles usaram DNA do óperon ao qual o repressor estava ligado e o trataram com a enzima DNase, que quebra o DNA. Eles foram capazes de recuperar

curtos filamentos de DNA que tinham sido protegidos da atividade enzimática pela molécula repressora e que supostamente constituíam a sequência do operador. A sequência de bases de cada filamento foi determinada, e cada mutação no operador foi mostrada como sendo uma mudança na sequência (Figura 10.12). Esses resultados mostraram que o locus operador é uma sequência específica de 17 a 25 nucleotídeos logo antes (a 5') do gene estrutural *Z*. Eles também mostraram a incrível especificidade de reconhecimento do repressor-operador, que pode ser perturbada por uma única substituição de base. Quando a sequência de bases no mRNA *lac* (transcrita do óperon *lac*) foi determinada, as primeiras 21 bases da ponta de iniciação 5' provaram ser complementares à sequência do operador que Gilbert havia determinado, mostrando que a sequência operadora é transcrita.

Os resultados desses experimentos deram uma confirmação crucial do mecanismo de ação repressora formulada por Jacob e Monod.

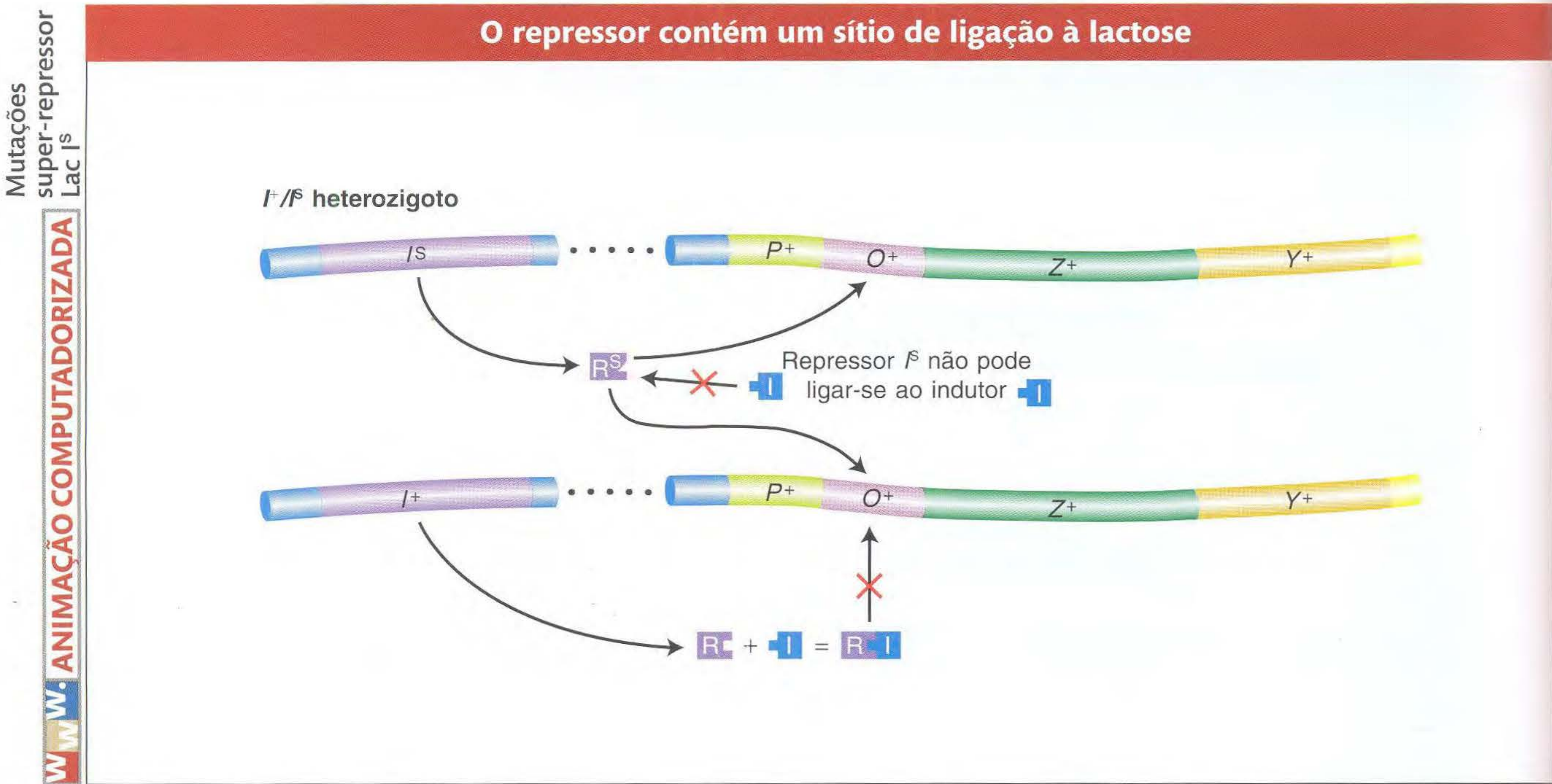


FIG. 10.10 A dominância da mutação I^s é devida à inativação do sítio alostérico do repressor Lac. Em uma célula diplóide I^s/I^- , nenhum dos genes estruturais *lac* são transcritos. O repressor I^s não tem um sítio de ligação de lactose (sítio alostérico) e, assim, não é inativado por um indutor. Portanto, mesmo na presença de um indutor, o repressor I^s liga-se irreversivelmente a todos os operadores em uma célula, bloqueando assim a transcrição do óperon *lac*.

A RNA polimerase faz contato com o promotor em seqüências específicas

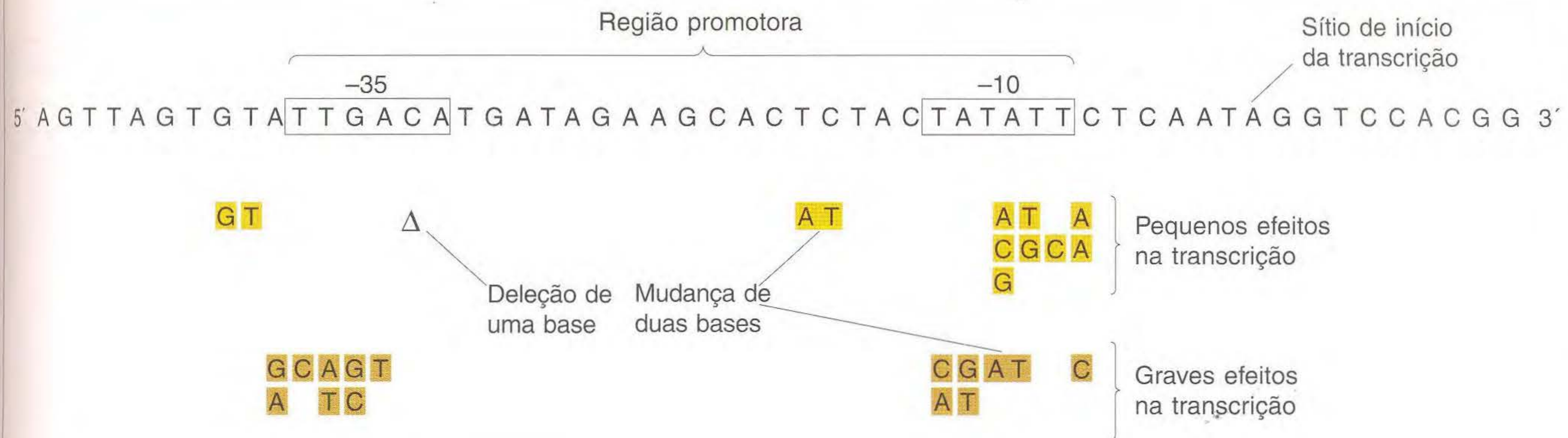


FIG. 10.11 Seqüências específicas de DNA são importantes para a transcrição eficiente dos genes de *E. coli* pela RNA polimerase. As seqüências em boxes são altamente conservadas em todos os promotores de *E. coli*, uma indicação de seu papel como sítios de contato na ligação ao DNA da RNA polimerase. As mutações nestas regiões têm um leve (amarelo) ou grave (marrom) efeito na transcrição. As mutações podem ser mudanças de nucleotídeos únicos ou pares de nucleotídeos, ou pode ocorrer uma deleção (Δ). [De J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller, *Recombinant DNA*, 2nd ed. Copyright 1992 by James D. Watson, Michael Gilman, Jan Witkowski, and Mark Zoller.]

O operador é uma seqüência específica de DNA

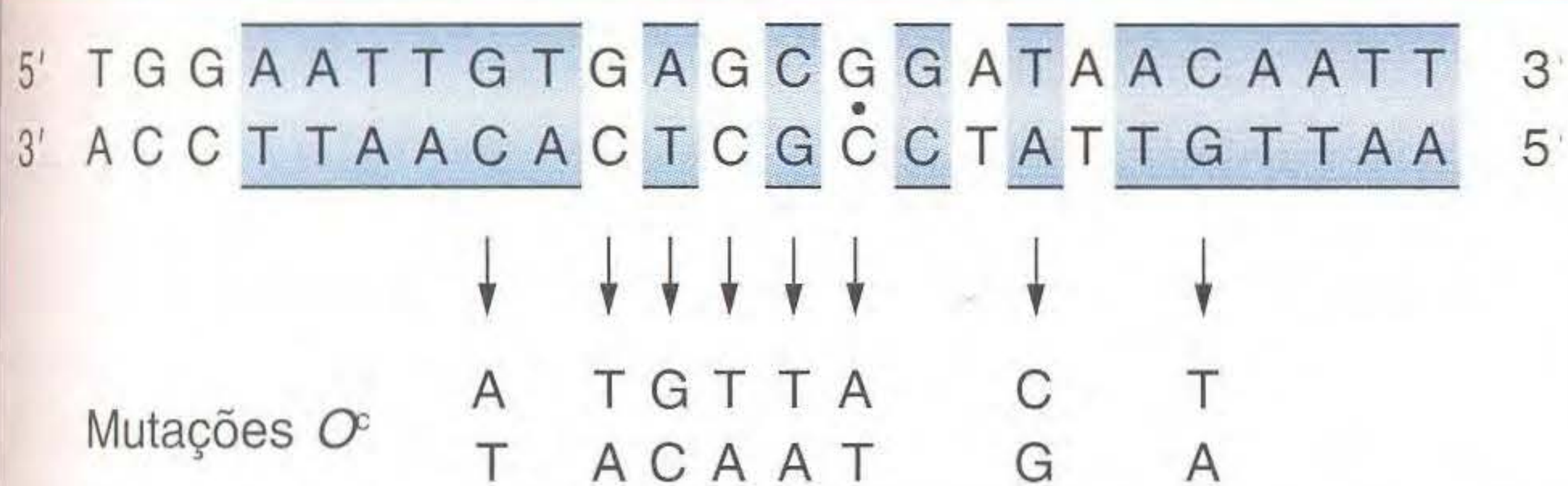


FIG. 10.12 A seqüência de bases do DNA do operador de lactose e as mudanças de bases associadas a oito mutações O^c . As regiões de simetria rotacional dupla são indicadas por cor e por um ponto em seu eixo de simetria. [De W. Gilbert, A. Maxam, and A. Mirzabekov, in N.O. Kjeldgaard and O. Malløe, Eds., *Control of Ribosome Synthesis*. Academic Press, 1976. Usada com permissão de Munksgaard International Publishers, Ltd., Copenhagen.]

Mutações polares

Algumas das mutações que foram mapeadas nos genes *Z* e *Y* foram vistas sendo polares, isto é, afetam genes “posteriores” ao óperon. Por exemplo, as mutações polares *Z* resultaram em um funcionamento nulo não só de *Z*, mas também de *Y* e *A*. As mutações polares em *Y* também afetaram *A*, mas não *Z*. Essas mutações polares foram as observações genéticas que sugeriram a Jacob e Monod que os três genes eram transcritos de uma ponta como uma unidade. As mutações polares resultaram de códons de fim que faziam com que os ribossomos saíssem do transcrito. Isso deixava um trecho nu de mRNA que era degradado, inativando assim os genes posteriores. (Os códons normais de fim e começo que fazem com que os ribossomos saiam e entrem no mRNA entre os genes estruturais não disparam essa degradação.)

10.3 Repressão Catabólica do Óperon *lac*: Controle Positivo

O sistema *lac* existente é aquele que, por um longo processo evolutivo, foi selecionado para operar de modo ótimo para

a eficiência de energia da célula bacteriana. Supostamente, para maximizar a eficiência energética, duas condições ambientais devem ser satisfeitas para que as enzimas do metabolismo de lactose sejam expressas.

Uma condição é que a lactose deve estar presente no ambiente. Essa condição faz sentido, pois seria ineficiente para a célula produzir as enzimas metabólicas de lactose se não houvesse substrato a ser metabolizado. Já vimos que o reconhecimento pela célula da presença de lactose é feito por uma proteína repressora.

A outra condição é que a glicose não pode estar presente no ambiente da célula. Como a célula pode captar mais energia da quebra da glicose do que da quebra de outros açúcares, é mais eficiente para a célula metabolizar glicose do que lactose. Assim, desenvolveram-se os mecanismos que impedem que a célula produza as enzimas para o metabolismo de lactose quando tanto lactose e glicose estiverem presentes juntas. A repressão da transcrição dos genes que metabolizam lactose na presença de glicose é um exemplo de **repressão catabólica** (a glicose é o produto da quebra, ou um *catabólito*, de lactose). A transcrição de genes que codificam proteínas necessárias para o metabolismo de muitos açúcares diferentes é similarmente reprimida na presença de glicose. Veremos que a repressão catabólica funciona por uma *proteína ativadora*.

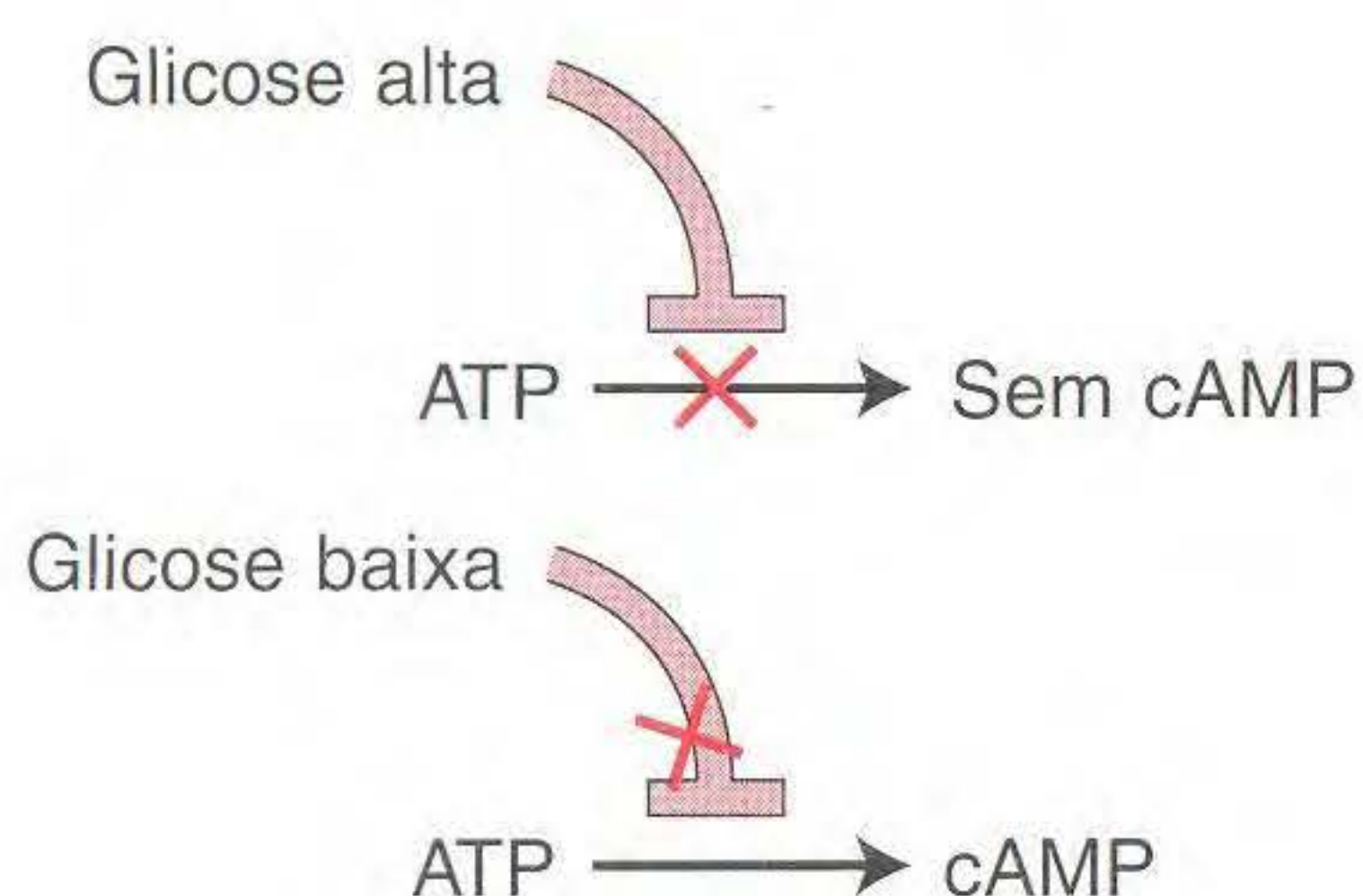
A base da repressão catabólica do óperon *lac*: escolher o melhor açúcar para metabolizar

Se tanto lactose quanto glicose estiverem presentes, a síntese de β -galactosidase não é induzida até que toda a glicose tenha sido usada. Assim, a célula conserva sua energia metabolizando qualquer glicose existente antes de passar pelo processo caro de criar uma nova maquinaria para metabolizar lactose.

Os resultados dos estudos indicam que um catabólito de glicose impede a ativação do óperon *lac* pela lactose, a repres-

Níveis de glicose controlam o óperon *lac*

(a) Níveis de glicose regulam níveis de cAMP



(b) Complexo cAMP-CAP ativa transcrição

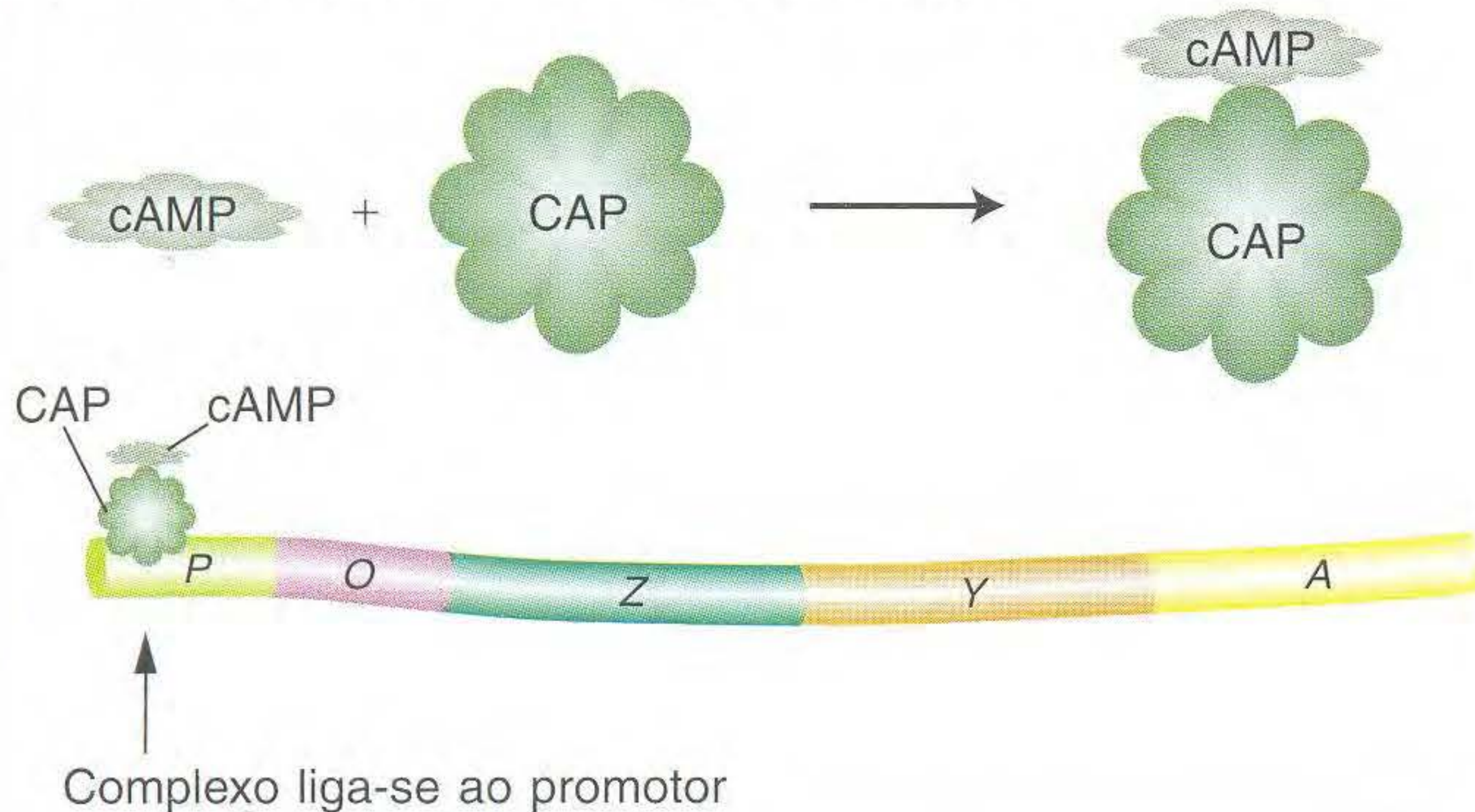


FIG. 10.13 Controle catabólico do óperon *lac*. (a) Apenas sob condições de baixa glicose é que é formado o cAMP (adenosina monofosfato cíclica). (b) Quando cAMP está presente, ela forma um complexo com CAP (proteína ativadora do catabolismo) que ativa a transcrição ligando-se a uma região dentro do promotor *lac*.

são catabólica mencionada. A identidade desse catabólito ainda é desconhecida. Entretanto, o catabólito de glicose é conhecido como modulando o nível de um importante constituinte celular, a **adenosina monofosfato cíclica (cAMP)**. Quando a glicose está presente em altas concentrações, a concentração de cAMP da célula é baixa. À medida que a concentração de glicose diminui, a concentração celular de cAMP aumenta correspondentemente. Uma alta concentração de cAMP é necessária para a ativação do óperon *lac*. Os mutantes que não podem converter ATP em cAMP não podem ser induzidos a produzir β -galactosidase, pois a concentração de cAMP não é grande o suficiente para ativar o óperon *lac*.

Qual é o papel do cAMP na ativação de *lac*? Um estudo de um conjunto diferente de mutantes deu a resposta. Esses mutantes fazem cAMP, mas não ativam as enzimas Lac, porque não têm uma outra proteína, chamada **proteína ativadora do catabolismo (CAP)**, codificada pelo gene *crp*. CAP liga-se a uma sequência específica de DNA do óperon *lac* (o sítio de ligação a CAP; veja Figura 10.14b). O DNA ligado a CAP é então capaz de interagir fisicamente com a RNA polimerase e aumentar essa afinidade enzimática pelo promotor *lac*. Por si mesma, CAP não pode ligar-se ao sítio de ligação a CAP do óperon *lac*. Entretanto, ligando-se a cAMP, seu efector alostérico, CAP é capaz de se ligar ao sítio de ligação a CAP e ativar a transcrição pela RNA polimerase. Pela inibição de CAP, quando a glicose está disponível, o sistema de repres-

são catabólica garante que o óperon *lac* será ativado apenas quando a glicose é pouca (Figura 10.13).

O óperon *lac* tem um nível adicional de controle, de modo que o óperon é inativo na presença de glicose mesmo se lactose também estiver presente. Um efector alostérico, cAMP, liga-se ao ativador CAP para permitir a indução do óperon *lac*. Entretanto, altas concentrações de catabólitos de glicose produzem baixas concentrações de cAMP, deixando de produzir cAMP-CAP e, assim, deixando de ativar o óperon *lac*.

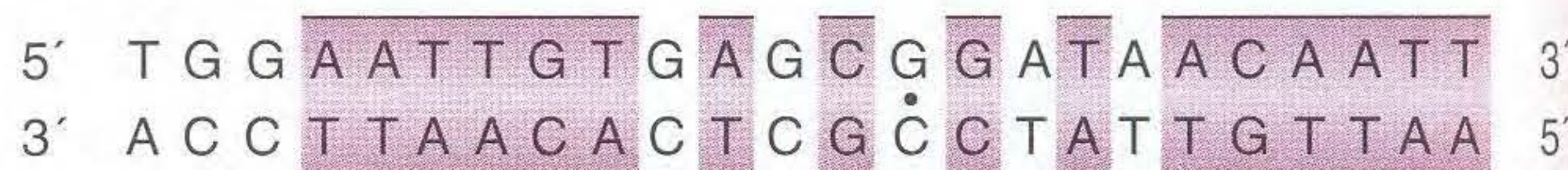
As estruturas dos sítios-alvo de DNA

As sequências de DNA às quais se liga o complexo CAP-cAMP hoje são conhecidas. Essas sequências (Figura 10.14) são muito diferentes das sequências às quais se liga o repressor Lac, embora ambas se liguem à ponta 5' do óperon. Essas diferenças destacam a especificidade da ligação ao DNA por essas proteínas reguladoras muito diferentes. Uma propriedade que essas sequências têm em comum e que é também comum a muitos outros sítios de ligação ao DNA é uma dupla simetria rotacional. Em outras palavras, se girarmos a sequência de DNA mostrada na Figura 10.14 por 180° dentro do plano da página, a sequência das bases destacadas dos sítios de ligação será idêntica. As bases destacadas são tidas como constituindo os sítios importantes de contato para as interações proteína-DNA. Essa simetria rotacional corresponde a simetrias dentro das proteínas de ligação ao DNA, muitas das quais são compostas de duas a quatro subunidades idênticas. Consideraremos as estruturas de algumas proteínas de ligação ao DNA mais adiante no capítulo.

Como a ligação do complexo cAMP-CAP ao operador aumenta a ligação da RNA polimerase ao promotor *lac*? Na Figura 10.15, o DNA é mostrado estando dobrado quando CAP se liga. Esse dobramento do DNA pode ajudar a ligação de RNA polimerase ao promotor. Também há evidências de que CAP faz contato direto com a RNA polimerase que é

Muitos sítios de ligação ao DNA são simétricos

(a) Operador *lac*



(b) Sítio de ligação a CAP

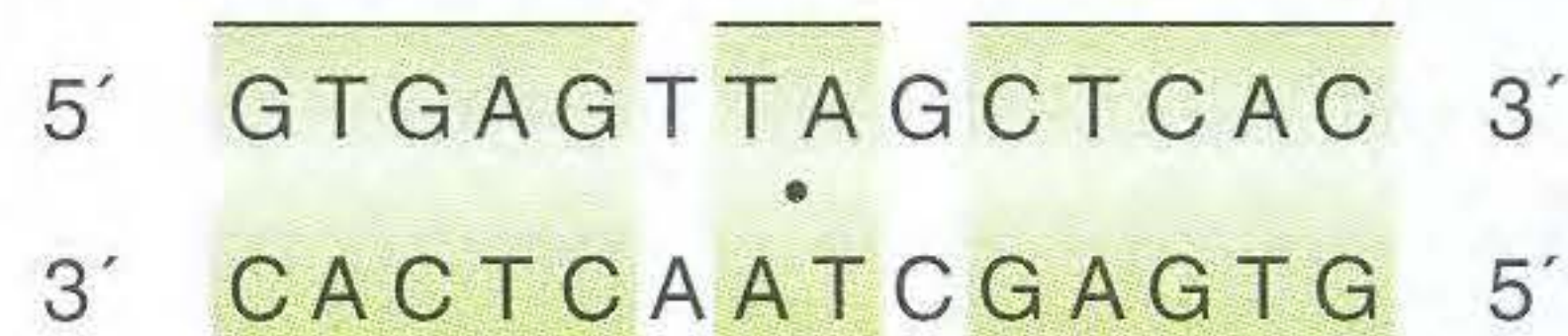
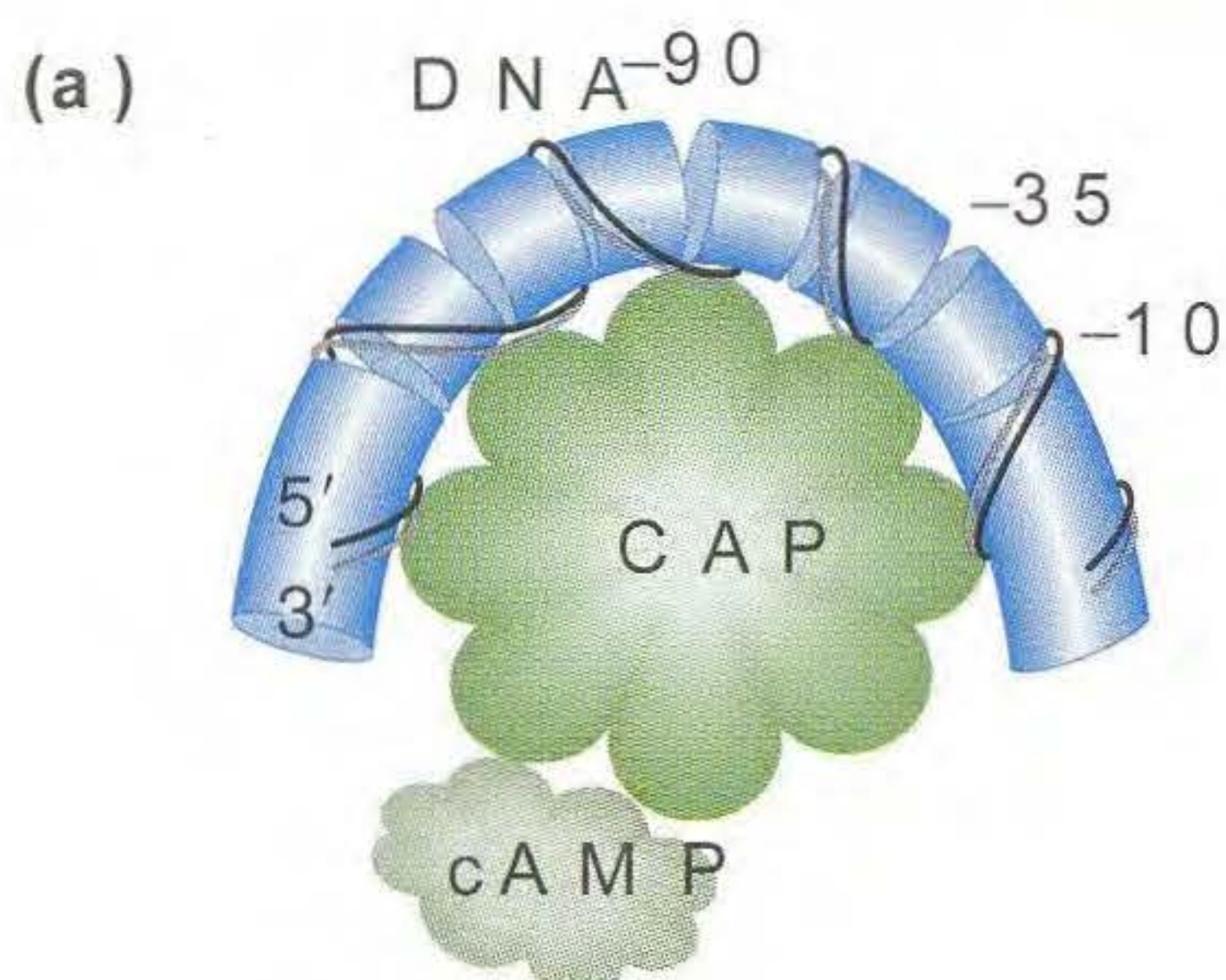


FIG. 10.14 A sequência de bases do DNA (a) do operador *lac*, ao qual se liga o repressor Lac, e (b) do sítio de ligação CAP, ao qual o complexo CAP-cAMP se liga. As sequências que apresentam simetria rotacional dupla são indicadas pelos boxes coloridos e por um ponto na parte central da simetria. [(a) De W. Gilbert, A. Maxam, and A. Mirzabekov, in N.O. Kjeldgaard and O. Malløe, Eds., *Control of Ribosome Synthesis*. Academic Press, 1976. Usada com permissão de Munksgaard International Publishers, Ltd., Copenhagen.]

A ligação de CAP dobra o DNA



(b)

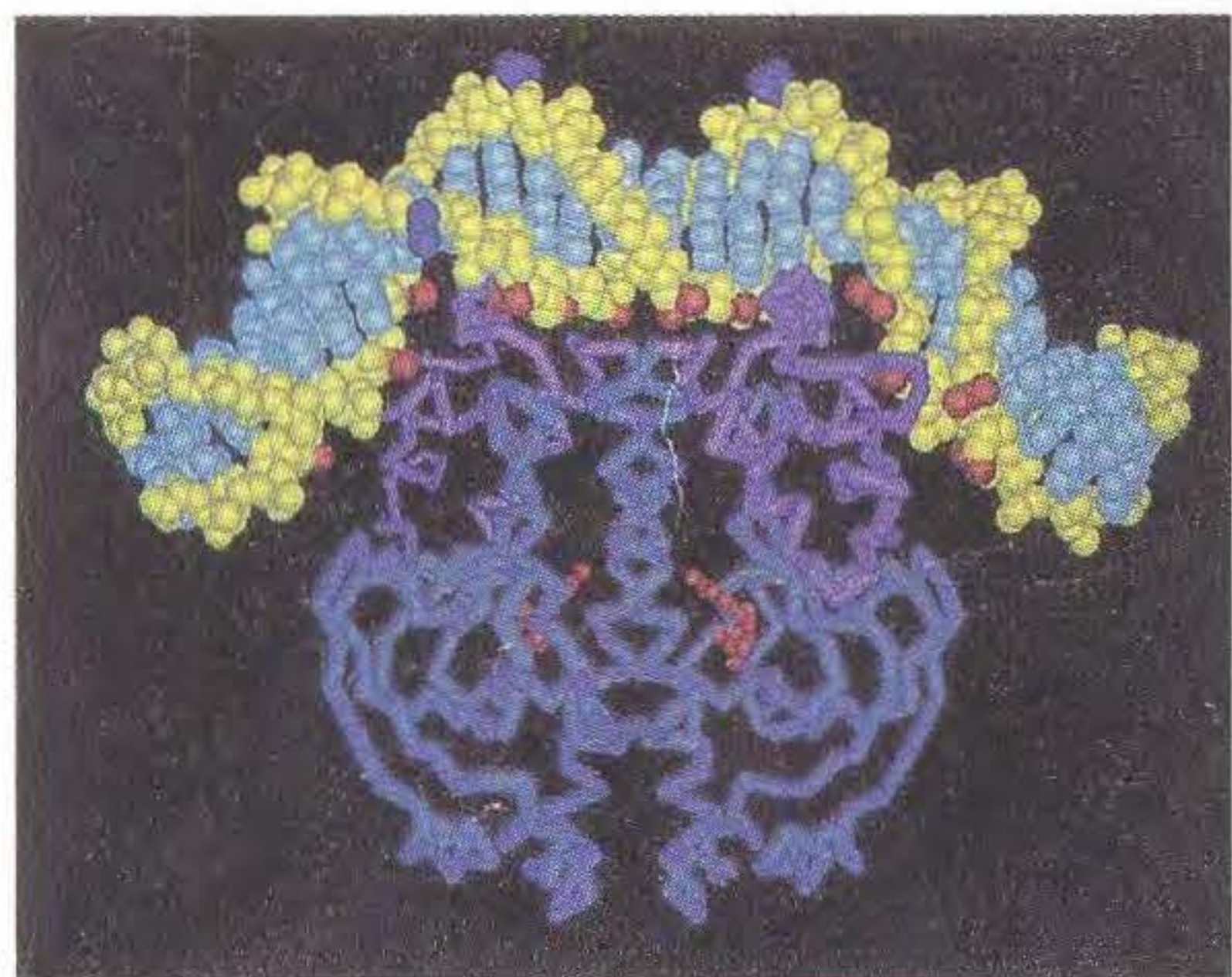


FIG. 10.15 (a) Quando CAP se liga ao promotor, cria uma dobra maior que 90° no DNA. (b) Imagem derivada da análise estrutural do complexo CAP-DNA. [(a) Redesenhado de B. Gartenberg and D.M. Crothers, *Nature* 333, 1988, 824. (Veja H.N. Lie-Johnson et al., *Cell* 47, 1986, 995.) De H. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, and J. Darnell, *Molecular Cell Biology*, 3rd ed. Copyright 1995 by Scientific American Books. (b) De L. Schultz and T.A. Steitz.]

importante para o efeito de ativação de CAP. A sequência de bases mostra que CAP e a RNA polimerase ligam-se diretamente adjacentes uma à outra no promotor *lac* (Figura 10.16).

Generalizando a partir do modelo de óperon *lac*, podemos imaginar o DNA como ocupado por proteínas reguladoras ligadas aos sítios operadores que elas controlam. O padrão exato de ligação irá depender de quais genes são ligados ou desligados e se os ativadores ou repressores regulam determinados óperons.

Um resumo do óperon *lac*

Agora podemos ajustar o CAP-cAMP e sítios de ligação da RNA polimerase ao modelo do óperon *lac*, como mostrado na Figura 10.17. A presença de glicose impede o metabolismo de lactose porque o produto da degradação de glicose inibe a manutenção dos altos níveis de cAMP necessários para a formação do complexo CAP-cAMP, que, por sua vez, é necessário para que a RNA polimerase ligue-se ao sítio promotor *lac*. Mesmo quando há poucos catabólitos de glicose e forma-se CAP-cAMP, o mecanismo para o metabolismo de lactose só será implementado se a lactose estiver presente. Esse nível de controle é obtido porque a lactose deve ligar-se à proteína repressora para removê-la do sítio operador e per-

mitir a transcrição do óperon *lac*. Assim, a célula conserva sua energia e recursos para produzir as enzimas que metabolizam a lactose apenas quando ambas são necessárias e úteis.

O controle indutor-repressor do óperon *lac* é um exemplo de repressão, ou **controle negativo**, no qual a expressão é normalmente bloqueada. Em contraste, o sistema CAP-cAMP é um exemplo de ativação, ou **controle positivo**, porque atua como um sinal que ativa a expressão — nesse caso, o sinal de ativação é a interação do complexo CAP-cAMP com o sítio de ligação de CAP ao DNA. A Figura 10.18 destaca esses dois tipos básicos de sistemas de controle.

O óperon *lac* é um aglomerado de genes estruturais que especificam enzimas que tomam parte do metabolismo de lactose. Esses genes são controlados pelas ações coordenadas das regiões do operador e promotor de ação cis. A atividade dessas regiões é, por sua vez, determinada pelas moléculas do repressor e ativador especificadas por genes reguladores.

► 10.4 Controle Duplo Positivo e Negativo: o Óperon de Arabinose

Como no sistema *lac*, o controle da transcrição em bactérias não é nem puramente positivo, nem puramente negativo; tanto a regulação positiva quanto a negativa podem controlar óperons individuais. A regulação do óperon de arabinose fornece um exemplo no qual uma única proteína de ligação ao DNA pode agir como um *repressor* ou um *ativador* (Figura 10.19), uma variação do tema geral de regulação transcricional por proteínas de ligação ao DNA.

Os genes estruturais *araB*, *araA* e *araD* codificam as enzimas metabólicas que degradam o açúcar arabinose. Os três genes são transcritos em uma unidade como um único mRNA. A transcrição é ativada em *araI*, a região **iniciadora**, que contém um sítio de ligação para uma proteína ativadora. O gene *araC*, que fica próximo, codifica uma proteína ativadora. Quando ligada a arabinose, essa proteína liga-se ao sítio *araI* e ativa a transcrição do óperon *ara*, talvez ajudando a RNA polimerase a se ligar ao promotor. Além disso, o mesmo sistema de repressão catabólica CAP-cAMP que impede a expressão do óperon *lac* na presença de glicose também impede a expressão do óperon *ara*.

Na presença de arabinose, tanto o complexo CAP-cAMP quanto o complexo AraC-arabinose devem ligar-se a *araI* para que a RNA polimerase se ligue ao promotor e transcreva o óperon *ara* (Figura 10.20a). Na ausência de arabinose, a proteína AraC adota uma conformação diferente e reprime o óperon *ara* ligando-se a *araI* e a um segundo sítio distante, *araO*, formando assim uma alça (Figura 10.20b) que impede a transcrição. Assim, a proteína AraC tem duas conformações, uma que atua como um ativador e outra que atua como repressor. A mudança liga/desliga do óperon é feita pela arabinose. As duas conformações, dependendo de se a arabinose efetora alostérica ligou-se a uma proteína, difere em suas habilidades de se ligar a um sítio-alvo específico na região *araO* do óperon.

CAP e RNA polimerase ligam-se próximos um do outro

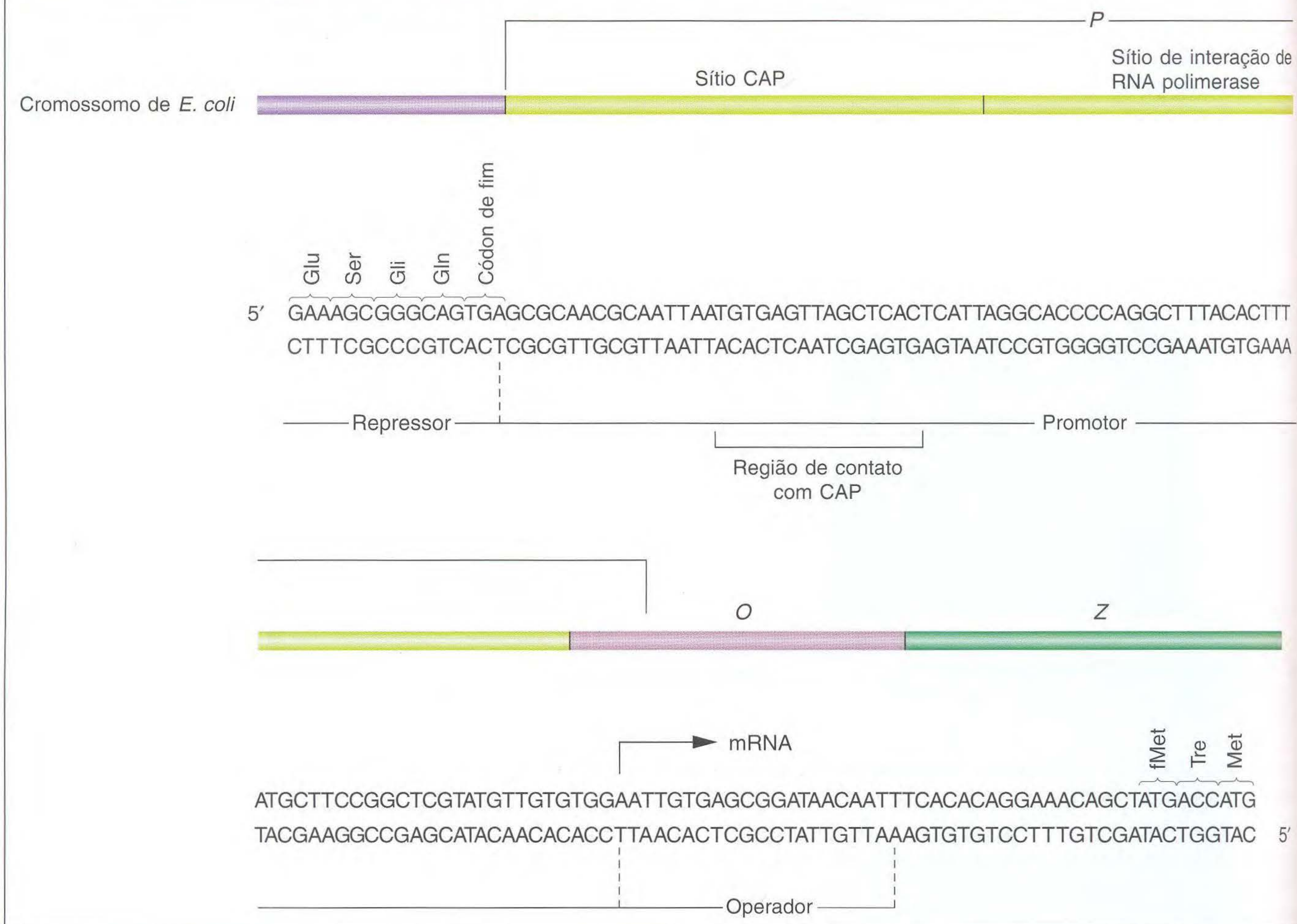


FIG. 10.16 A região controladora do operon *lac*. A sequência de base e os limites genéticos da região controladora do operon *lac*, com seqüências parciais para genes estruturais. [De R.C. Dickson, J. Abelson, W.M. Barnes, and W.S. Reznikoff, "Genetic Regulation: The Lac Control Region," *Science* 187, 1975, 27. Copyright 1975 by the American Association for the Advancement of Science.]

A transcrição do operon pode ser regulada tanto por ativação quanto por repressão. Os operons que regulam o metabolismo de compostos similares, tais como açúcares, podem ser regulados de modos bem diferentes.

► 10.5 Vias Metabólicas e Níveis Adicionais de Regulação: Atenuação

O controle coordenado de genes em bactérias é generalizado. Na seção precedente, vimos exemplos ilustrando as vias de regulação para a degradação de açúcares específicos. De fato, a maioria do funcionamento gênico coordenado em bactérias atua por meio de mecanismos de operon. Em muitas vias que sintetizam moléculas essenciais a partir de blocos estruturais inorgânicos, os genes que codificam as enzimas são organizados em operons, completados com mRNA multigênicos.

Além disso, nos casos em que é conhecida a atividade catalítica, há uma marcante congruência entre a ordem dos genes no operon no cromossomo e a ordem na qual seus produtos atuam na via metabólica. Essa congruência é bem ilustrada pela organização do operon de triptofano em *E. coli* (Figura 10.21). O operon de triptofano contém cinco genes (*trpE*, *trpD*, *trpC*, *trpB*, *trpA*) que codificam enzimas que contribuem para a síntese do aminoácido triptofano.

Em bactérias, os genes que codificam enzimas que estão na mesma via metabólica são geralmente organizados em operons.

A regulação do triptofano e alguns outros operons funcionando na biossíntese de aminoácidos envolve dois mecanismos de regulação transcricional. Um dá o controle global da expressão do mRNA do operon e outro fornece um controle fino.

Controle negativo e positivo do óperon *lac*

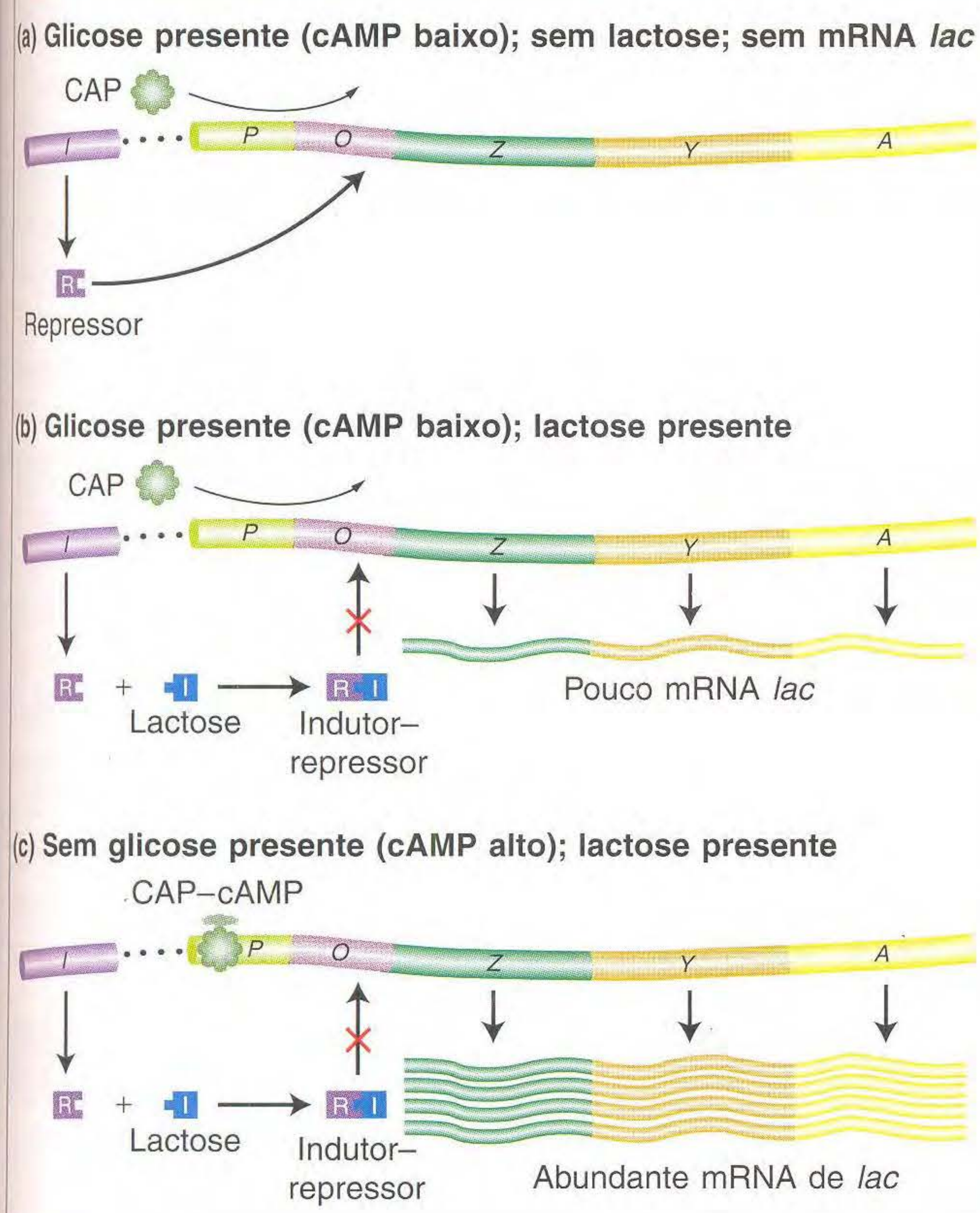


FIG. 10.17 O óperon *lac* é controlado conjuntamente pelo repressor Lac (controle negativo) e a proteína ativadora do catabolismo (CAP) (controle positivo). Grandes quantidades de mRNA são produzidas apenas quando a lactose está presente para inativar o repressor, e baixos níveis de glicose promovem a formação do complexo CAP-cAMP, que regulam positivamente a transcrição. [Redesenhado de B. Gartenberg and D.M. Crothers, *Nature* 333, 1988, 824. (Veja H.N. Lie-Johnson et al., *Cell* 47, 1986, 995.) De H. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, and J. Darnell, *Molecular Cell Biology*, 3rd ed. Copyright 1995 by Scientific American Books.]

A transcrição do óperon *trp* é regulada em duas etapas

O nível da expressão gênica do óperon *trp* é controlado pelo nível de triptofano. Quando o triptofano está ausente do meio de crescimento, a expressão do gene *trp* é alta; quando os níveis de triptofano são altos, o óperon *trp* é reprimido. Um mecanismo para o controle fisiológico da transcrição do óperon *trp* emprega o repressor Trp, o produto do gene *trpR*. O repressor Trp liga triptofano e desliga a transcrição do óperon quando os níveis adequados de aminoácidos estão presentes. Esse mecanismo simples garante que a célula não gaste energia, produzindo triptofano, quando o aminoácido é suficientemente abundante. As linhagens de *E. coli* com mutações em *trpR* continuam a expressar o mRNA de *trp* e, assim, continuam a produzir triptofano quando o aminoácido é abundante.

Ao estudar essas linhagens mutantes de *trpR*, Charles Yanofsky descobriu que, quando o triptofano era removido do meio, a produção de mRNA de *trp* aumentava várias vezes. Esse achado foi a evidência de que um segundo mecanismo

Repressão e ativação comparadas

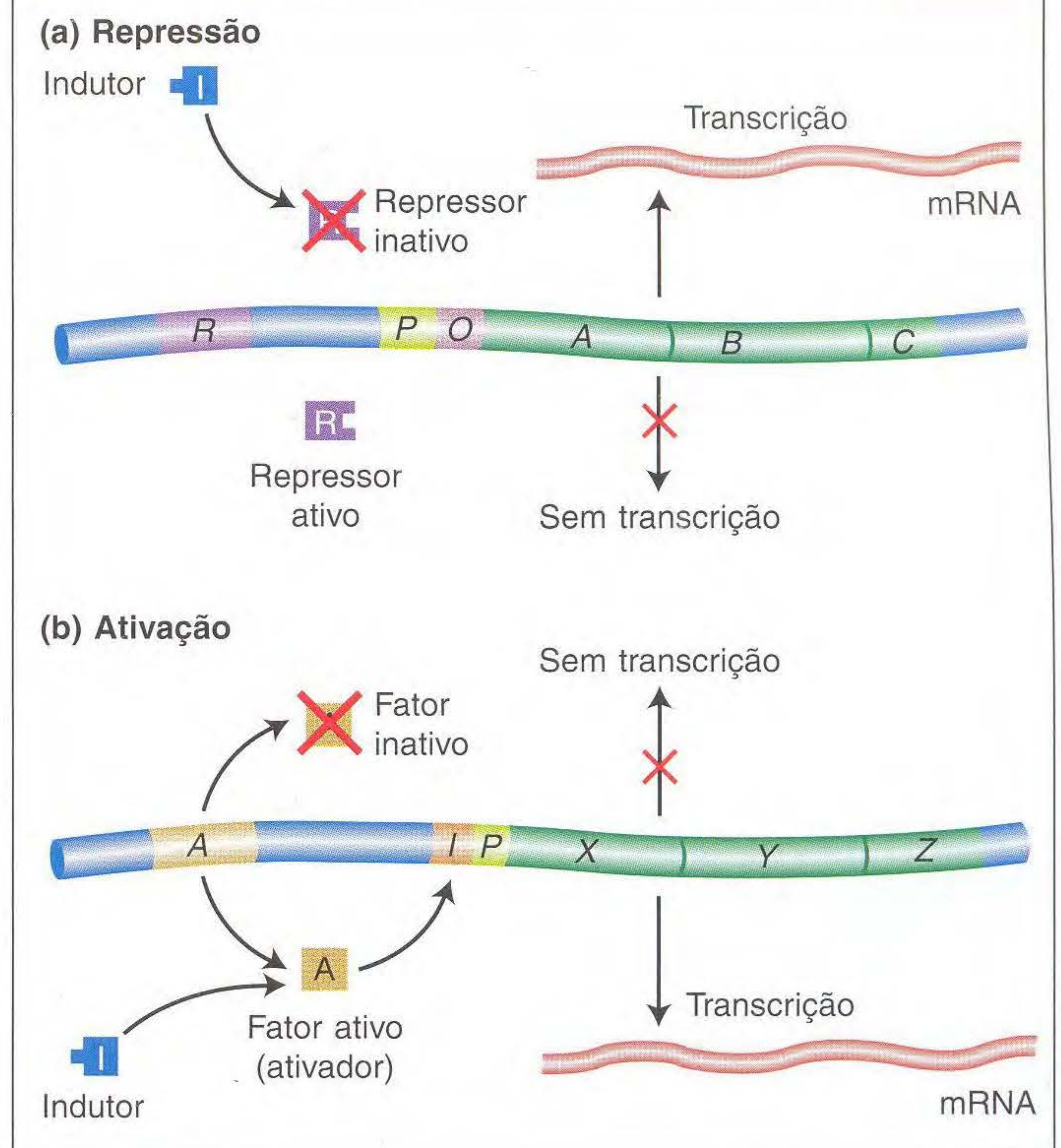


FIG. 10.18 (a) Na repressão, um repressor ativo (codificado pelo gene *R* nesse exemplo) bloqueia a expressão do óperon *A, B, C* ligando-se a um sítio operador (*O*). (b) Na ativação, um ativador funcional é necessário para a expressão gênica. Um ativador não-funcional resulta na falta de expressão dos genes *X, Y, Z*. Pequenas moléculas podem converter um ativador não-funcional em um funcional, que então se liga à região controladora do óperon, chamada de *I* nesse caso. As posições de ambos, *O* e *I*, com relação ao promotor *P* nos dois exemplos estão arbitrariamente desenhadas, pois suas posições diferem em óperons diferentes.

de controle existia para regular negativamente a transcrição além do repressor Trp. Esse mecanismo é chamado de **atenuação** porque a produção de mRNA é normalmente *atenuada*, isto é, “diminuída”, quando o triptofano é abundante. Ao contrário de outros mecanismos de controle bacteriano descritos até agora, a atenuação atua uma etapa *após* o início da transcrição.

Os mecanismos que controlam a atenuação foram descobertos identificando mutações que reduziam ou aboliam a atenuação. As linhagens com essas mutações produziam mRNA de *trp* em níveis máximos mesmo na presença de

Mapa do óperon *ara*

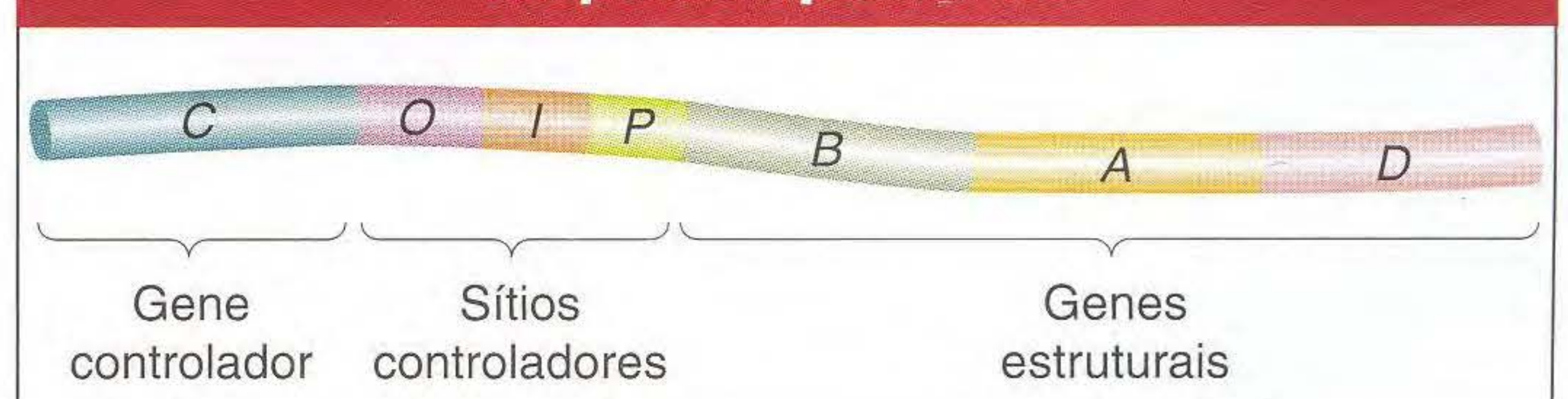


FIG. 10.19 Os genes *B, A* e *D* juntos com sítios *I* e *O* constituem o óperon *ara*. *O* é *araO* e *I* é *araI*.

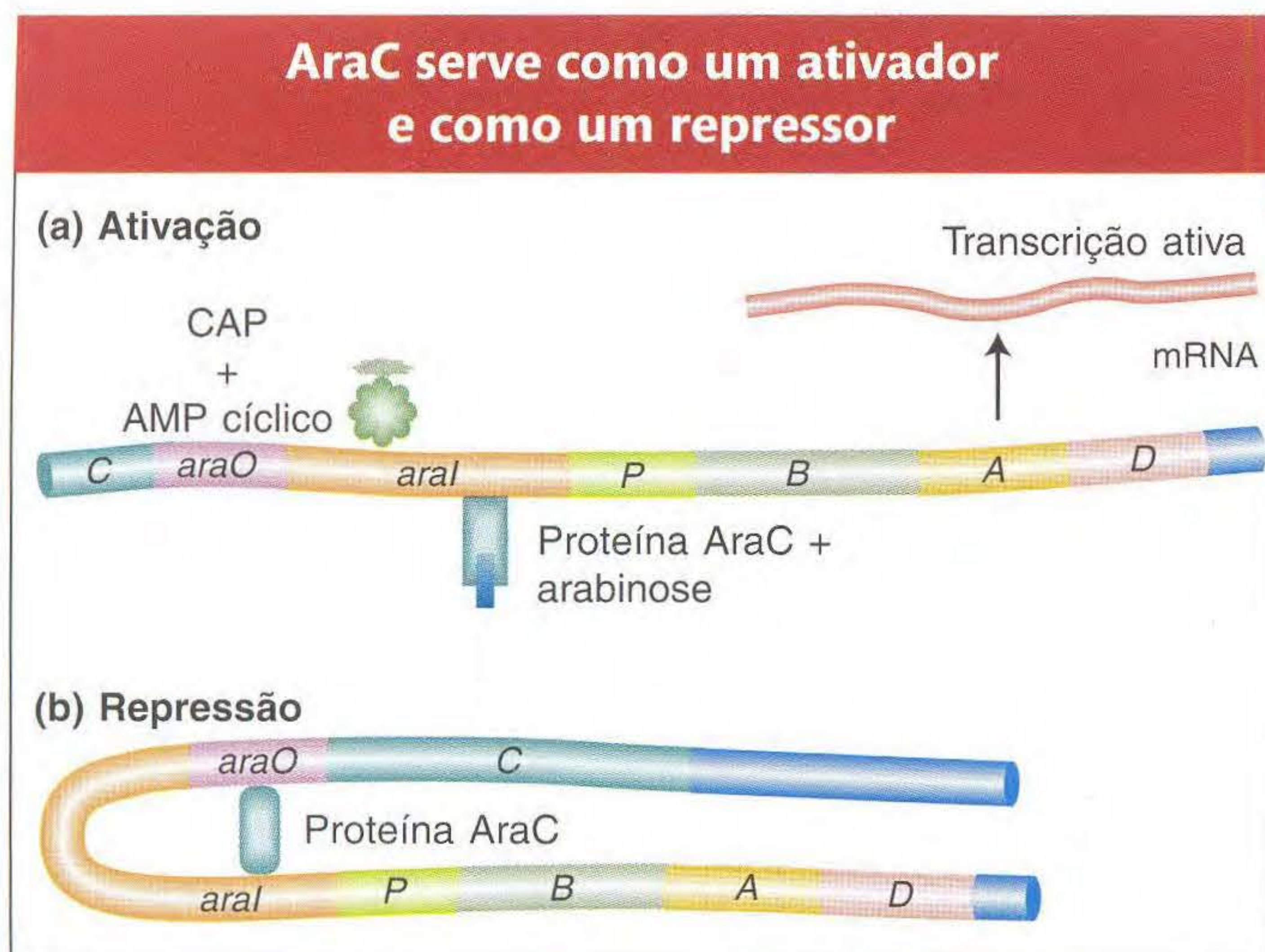


FIG. 10.20 Controle duplo do operon *ara*. (a) Na presença de arabinose, a proteína AraC liga-se à região *araI*. O complexo CAP-cAMP liga-se a um sítio adjacente a *araI*. Essa ligação estimula a transcrição dos genes *araB*, *araA* e *araD*. (b) Na ausência de arabinose, a proteína AraC liga-se tanto à região *araI* quanto à *araO*, formando uma alça de DNA. Essa ligação impede a transcrição do operon *ara*.

triptofano. Yanofsky mapeou as mutações em determinada região entre o operador *trp* e o gene *trpE*. Essa região, chamada de **seqüência líder**, está na ponta 5' do mRNA do operon *trp* antes do primeiro códon do gene *trpE* (Figura 10.22). A seqüência líder de *trp* é incomumente maior para um mRNA procariótico, 160 bases, e a análise detalhada revelou como uma parte dessa seqüência funciona como um **atenuador** que controla a posterior transcrição do mRNA *trp*.

As observações principais são que, na ausência da proteína repressora TrpR, a presença de triptofano pára a transcrição após cerca das primeiras 140 bases, enquanto, na ausência de triptofano para a transcrição, a transcrição do operon con-

tinua. O mecanismo de término ou continuação compreende dois elementos-chave. Primeiro, a seqüência líder do mRNA *trp* codifica um curto peptídeo de 14 aminoácidos que inclui dois códons adjacentes de triptofano. O triptofano é um dos aminoácidos menos abundantes nas proteínas, e é codificado por um único códon. Esse par de códons de triptofano é, portanto, uma característica incomum. Segundo, partes do líder do mRNA de *trp* formam estruturas haste e alça que são capazes de alternar entre duas conformações. Uma dessas conformações favorece o término da transcrição (Figura 10.23a).

A lógica regulatória do operon está na abundância de triptofano. Quando o triptofano é abundante, há um fornecimento suficiente de tRNA^{Trp} para permitir a tradução do peptídeo de 14 aminoácidos. Lembre que a transcrição e tradução em bactérias são acopladas; logo, os ribossomos podem se ligar aos mRNA transcritos e iniciar as traduções antes que a transcrição esteja completa. A ligação do ribossomo altera as estruturas secundárias do mRNA de *trp* de tal modo que a transcrição é terminada (Figura 10.23b). Entretanto, quando o triptofano está limitado, o ribossomo é parado nos códons do triptofano e a transcrição é capaz de continuar (Figura 10.23c).

Outros operons para enzimas nas vias biossintéticas têm controles de atenuação similares. Uma marca dos operons de biossíntese de aminoácidos é a presença de vários códons para determinado aminoácido em um peptídeo separado codificado pela seqüência líder 5'. Por exemplo, o operon de *fen* tem sete códons de fenilalanina em um peptídeo de 15 aminoácidos e o operon *his* tem sete códons de histidina em tandem em seu peptídeo líder (Figura 10.24.)

Um segundo nível de regulação nos operons de biossíntese de aminoácidos é a atenuação da transcrição mediada pela abundância do aminoácido e a tradução de um peptídeo líder.

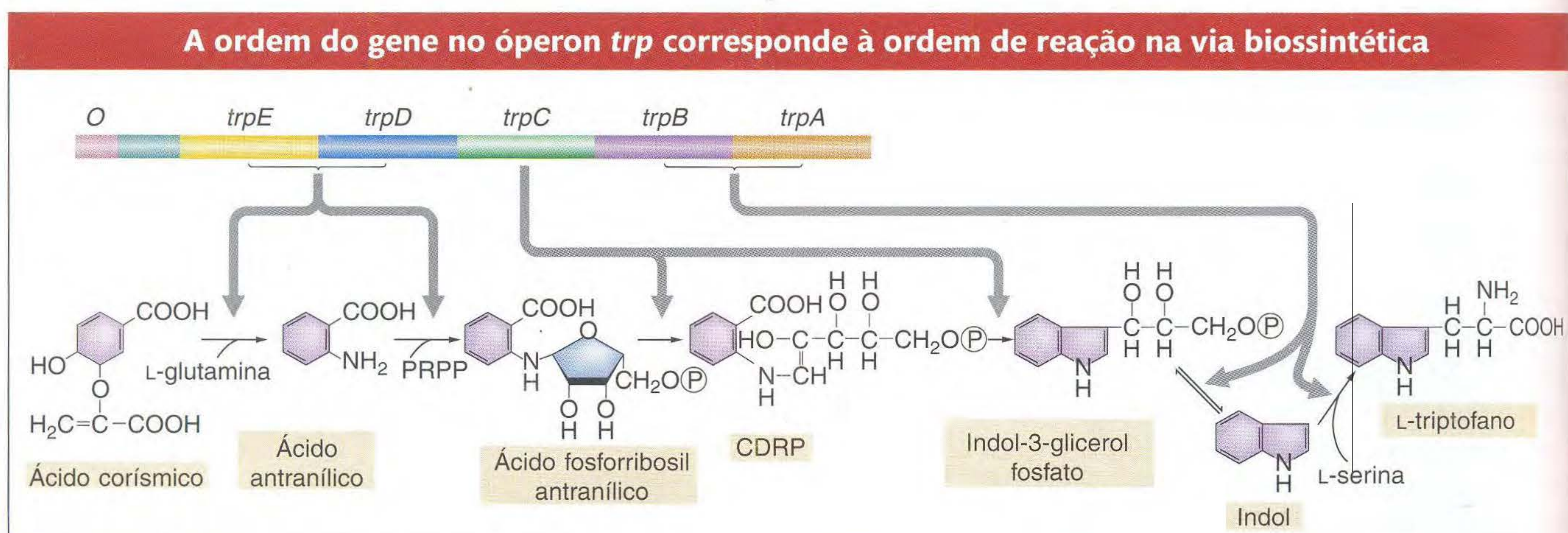


FIG. 10.21 A ordem cromossômica dos genes no operon *trp* de *E. coli* e a seqüência de reações catalisadas pelos produtos enzimáticos dos genes estruturais *trp*. Os produtos dos genes *trpD* e *trpE* formam um complexo que catalisa etapas específicas, como fazem os produtos dos genes *trpB* e *trpA*. A triptofano sintetase é uma enzima tetramérica formada pelos produtos de *trpB* e *trpA*. Ela catalisa um processo de duas etapas que leva à formação de triptofano. Abreviações: PRPP, fosforribosilpirofosfato; CDRP, 1-(o-carboxifenilamino)-1-desoxirribulose 5-fosfato. [De S. Tanemura and R.H. Bauerle, *Genetics* 95, 1980, 545.]

A sequência líder do mRNA de *trp* contém uma região atenuadora e dois códons de triptofano

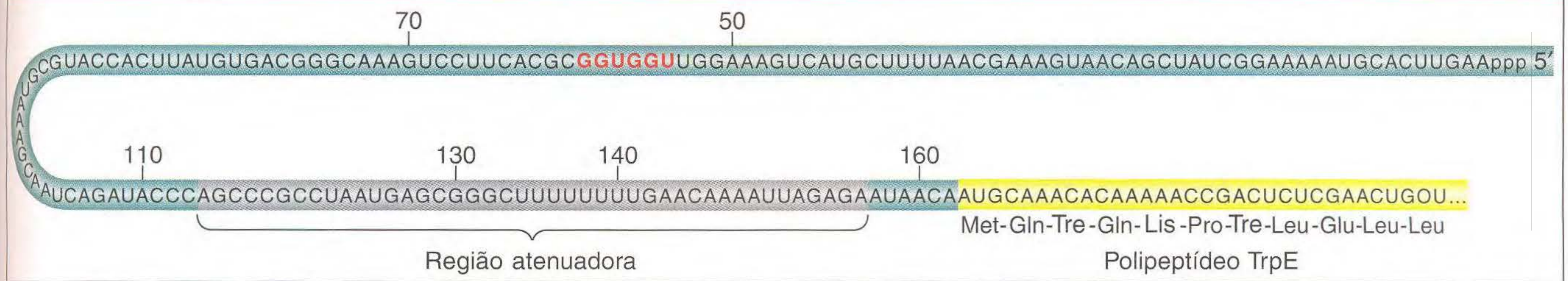


FIG. 10.22 Na sequência líder do mRNA de *trp*, a região atenuadora precede a sequência codificante *trpE*. Mais antecedente, nas bases 54 até 59, estão dois códons de triptofano (mostrados em vermelho) do peptídeo líder. [De G.S. Stent and R. Calendar, *Molecular Genetics*, 2nd ed. Copyright 1978 by W.H. Freeman and Company. Com base em dados não-publicados fornecidos por Charles Yanofsky.]

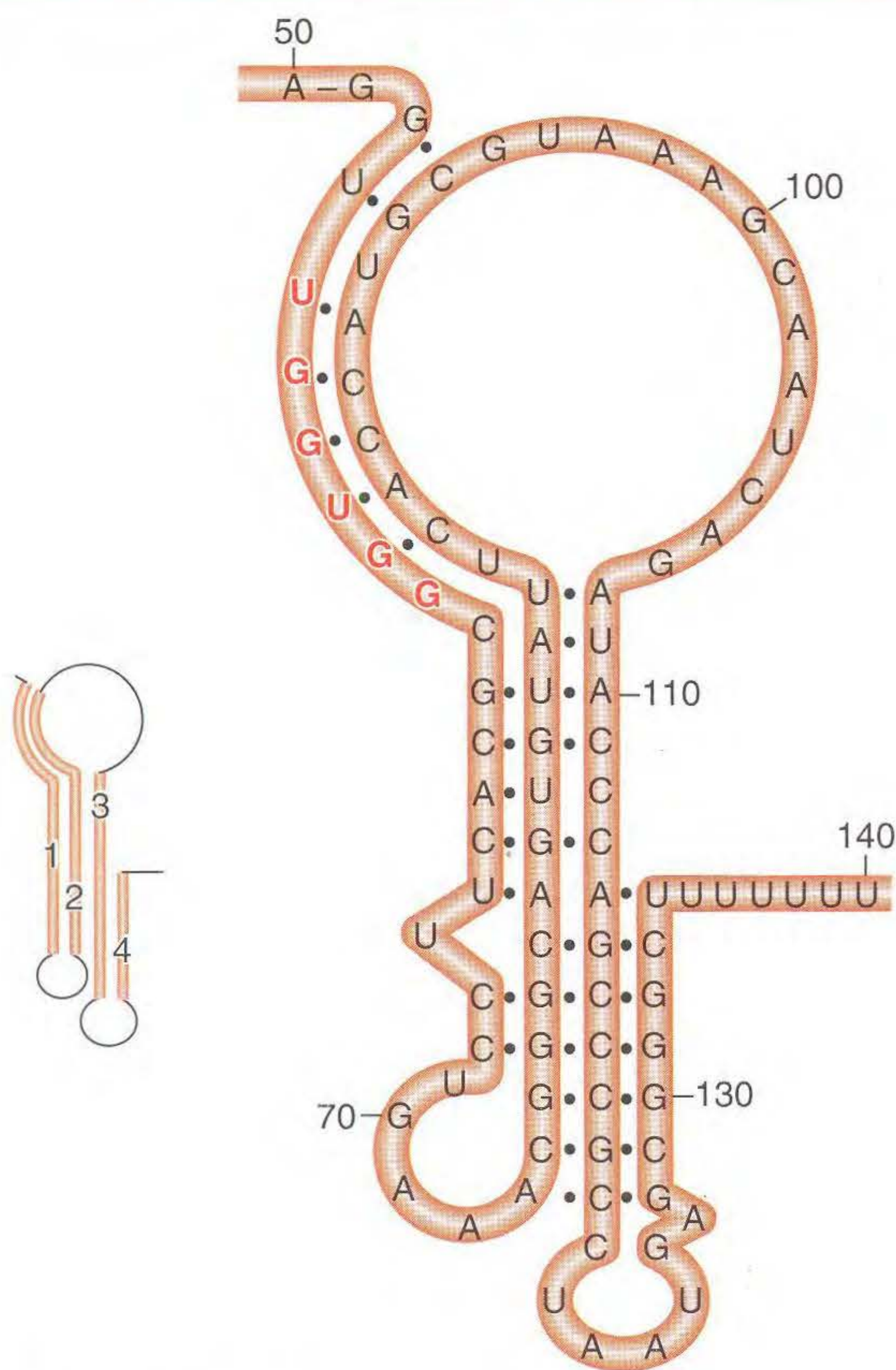
10.6 Ciclos de Vida de Bacteriófagos: Mais Reguladores, Óperons Complexos

Naquele cinema em Paris, François Jacob teve um *insight* de que o fenômeno da indução do profago poderia ser muito

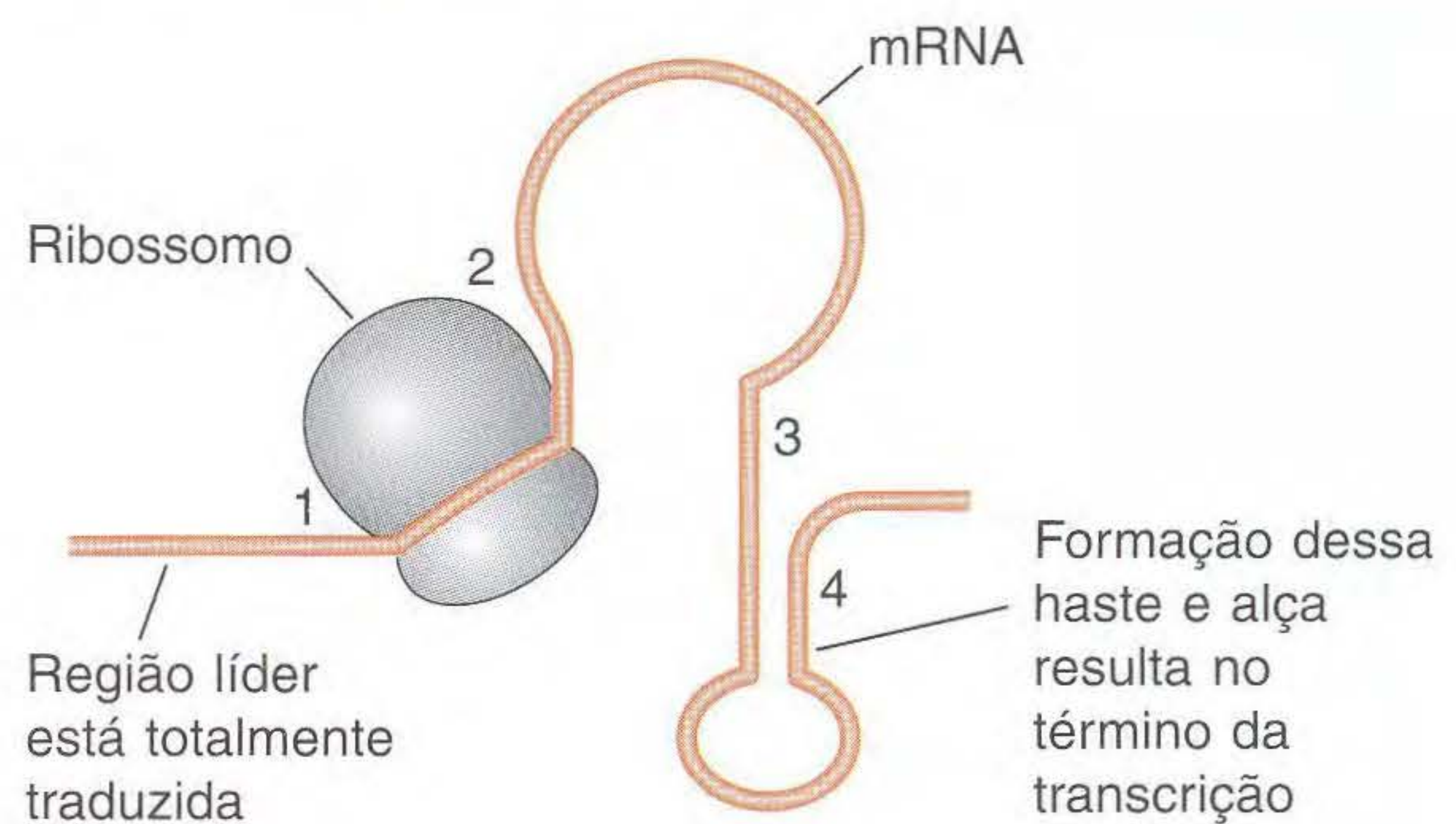
análogo à indução da síntese da enzima β -galactosidase. Ele estava certo. Aqui, veremos como o ciclo de vida e o genoma do bacteriófago λ são regulados. Embora sua regulação seja mais complexa que a dos óperons individuais, eles são controlados por modos agora familiares de regulação gênica.

O bacteriófago λ é um chamado fago temperado que tem dois ciclos de vida alternativos (Figura 10.25). Quando uma

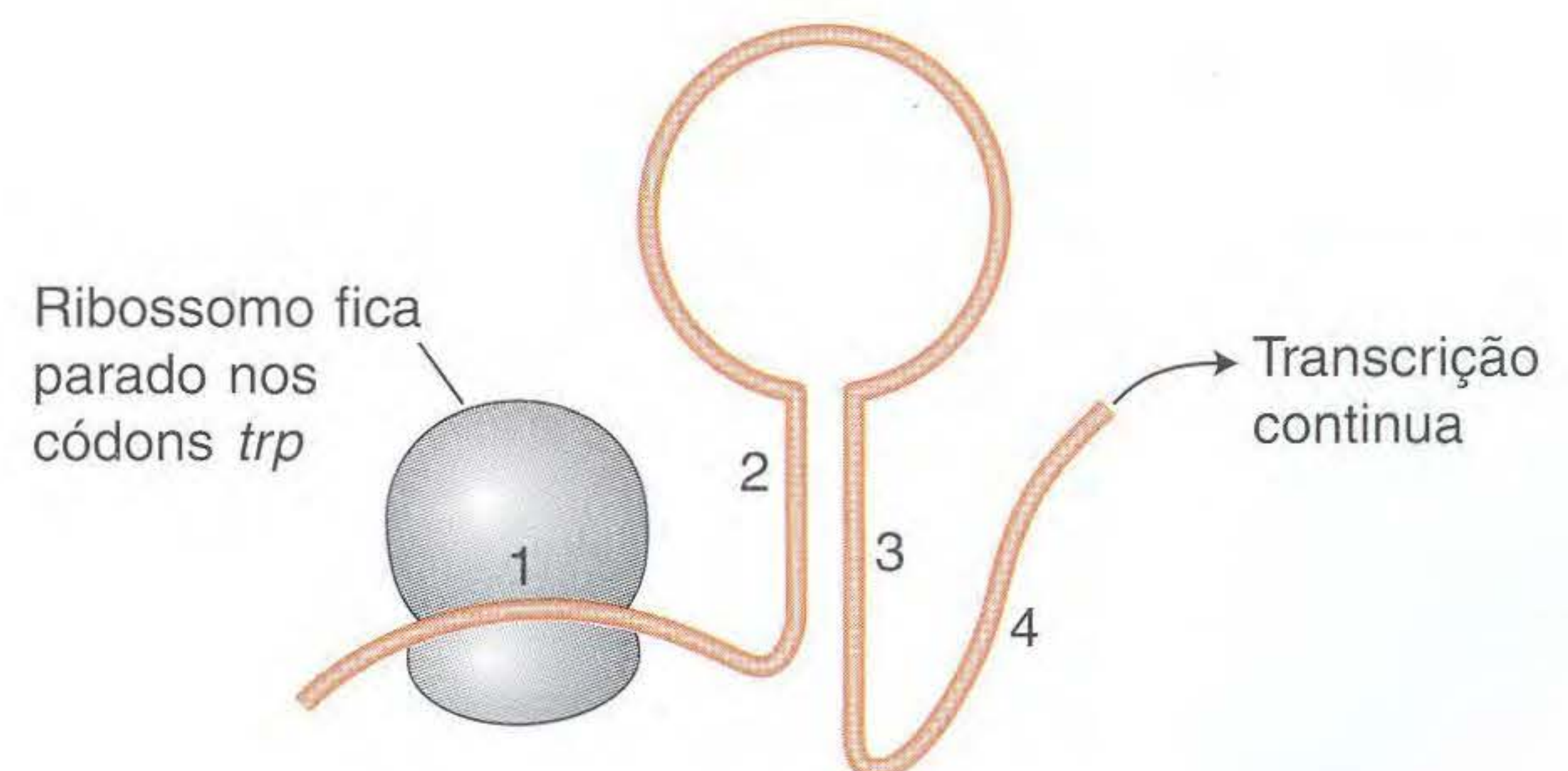
O triptofano abundante atenua a transcrição do óperon *trp*



(a) mRNA líder *trp*



(b) Alto nível de triptofano



(c) Nível baixo de triptofano

FIG. 10.23 Modelo para atenuação do óperon *trp*. (a) Estruturas secundárias propostas na conformação do mRNA líder *trp* que favorece o término da transcrição. Quatro regiões podem formar pares de bases que geram três estruturas haste-alça. (b) Quando o triptofano é abundante, o segmento 1 do mRNA de *trp* é traduzido. O segmento 2 entra no ribossomo (embora não seja traduzido), o que permite que os segmentos 3 e 4 façam pares de bases. Essa região de pares de bases faz com que a RNA polimerase termine a transcrição. (c) Em contraste, quando o triptofano é pouco, o ribossomo pára nos códons do segmento 1. O segmento 2 interage com o segmento 3 em vez de ser levado para o ribossomo, e, assim, os segmentos 3 e 4 não podem parear. Conseqüentemente, a transcrição continua. [De D.L. Oxender, G. Zurawski, and C. Yanofsky, *Proc. Natl. Acad. Sci. USA* 76, 1979, 5524.]

Peptídeos-líder de óperons de biossíntese de aminoácidos

(a) Óperon <i>trp</i>	Met - Lis - Ala - Ile - Fen - Val - Leu - Lis - Gli - Trp - Trp - Arg - Tre - Ser - Fim	
5'	AUG - AAA - GCA - AUU - UUC - GUA - CUG - AAA - GGU - UGG - UGG - CGC - ACU - UCC - UGA	3'
(b) Óperon <i>fen</i>	Met - Lis - His - Ile - Pro - Fen - Fen - Fen - Ala - Fen - Fen - Fen - Tre - Fen - Pro - Fim	
5'	AUG - AAA - CAC - AUA - CCG - UUU - UUU - UUC - GCA - UUC - UUU - UUU - ACC - UUC - CCC - UGA	3'
(c) Óperon <i>his</i>	Met - Tre - Arg - Val - Gln - Fen - Lis - His - His - His - His - His - His - His - Pro - Asp	
5'	AUG - ACA - CGC - GUU - CAA - UUU - AAA - CAC - CAC - CAU - CAU - CAC - CAU - CAU - CCU - GAC	3'

FIG. 10.24 (a) A parte traduzida de (a) da região líder *trp* contém dois códons consecutivos de triptofano, (b) a sequência líder *fen* contém sete códons de fenilalanina e (c) a sequência líder *his* contém sete códons consecutivos de histidina.

bactéria normal é infectada por um fago λ tipo selvagem, podem ocorrer dois resultados possíveis: (1) o fago pode se replicar e, eventualmente, lisar a célula (o **ciclo lítico**) ou (2) o genoma do fago pode se integrar ao cromossomo da bactéria como um profago inerte (o **ciclo lisogênico**). No estado lítico, a maioria dos 71 genes do fago são expressos em algum ponto, enquanto, no estado lisogênico, a maioria dos genes estão inativos.

Como decidir qual das duas vias seguir? O controle fisiológico da decisão entre a via lítica ou lisogênica depende dos recursos disponíveis na bactéria hospedeira. Se os recursos forem abundantes, existem nutrientes suficientes para fazer muitas cópias do vírus e o ciclo lítico é o preferido. Se os recursos são limitados ou na margem, então a via lisogênica é escolhida. O vírus então permanece como um *profago* até que as condições melhorem ou mudem. O profago inerte pode ser induzido por luz ultravioleta a entrar no ciclo lítico, o fenômeno estudado por Jacob. Os estados lítico e lisogênico são caracterizados por programas de expressão gênica bem distintos que devem ser regulados. Que estado alternativo é selecionado é determinado por uma mudança genética complexa compreendendo várias proteínas reguladoras de ligação ao DNA e um conjunto de sítios operadores.

Do mesmo modo que para os sistemas reguladores de *lac* e outros, as análises genéticas dos mutantes foram fontes de *insights* cruciais dos componentes e lógica da mudança genética λ . As triagens fenotípicas simples possibilitaram o isolamento dos mutantes que foram defectivos tanto na via lítica quanto na lisogênica. Os mutantes de cada tipo podiam ser reconhecidos pelo aspecto das placas infectadas em uma camada de bactérias. Quando as partículas de fago tipo selvagem eram colocadas em uma camada de bactérias sensíveis, apareciam uns claros (placas) onde as bactérias tinham sido infectadas e lisadas, mas essas placas são turvas porque as bactérias que foram lisogenizadas cresceram dentro delas. Os fagos mutantes que formam placas claras podem assim ser selecionados, pois esses fagos são incapazes de lisogenizar as células.

Tais mutantes *claros* (designados por *c*) são análogos aos mutantes *I* e *O* do sistema *lac*. Esses mutantes foram isolados

como mutantes sensíveis à temperatura que têm fenótipos *claros* em temperaturas mais altas, mas fenótipos tipo selvagem em temperaturas mais baixas. Três classes de mutantes levaram à identificação de características reguladoras chave do fago λ . Na primeira classe, os mutantes para os genes *cI*, *cII* e *cIII* formam placas claras; isto é, eles são incapazes de estabelecer lisogenia. Foi isolada uma segunda classe de mutantes que não lisogenizam células, mas podem se replicar e entrar no ciclo lítico em uma célula lisogenizada. Esses mutantes são análogos aos mutantes com operador constitutivo do sistema *lac*. Um terceiro mutante-chave pode lisogenizar, mas é incapaz de lisar as células. O gene mutante nesse caso é o gene *cro* (de controle do repressor e outras coisas). A decisão entre as vias lítica e lisogênica depende da atividade das proteínas codificadas pelos quatro genes *cI*, *cII*, *cIII* e *cro*, três dos quais são proteínas de ligação ao DNA.

Primeiro enfocaremos dois genes *cI* e *cro* e as proteínas que eles codificam. O gene *cI* codifica um repressor, em geral chamado de repressor λ , que reprime o crescimento lítico e promove a lisogenia. O gene *cro* codifica um repressor que reprime a lisogenia, permitindo assim o crescimento lítico. A mudança genética que controla os dois ciclos de vida do fago λ tem dois estados: no lisogênico, *cI* está ligado, mas *cro* está desligado; e, no ciclo lítico, *cro* está ligado, mas *cI* está desligado. Portanto, o repressor λ e Cro estão em competição, e qual dos repressores prevalece determina o estado da mudança e a expressão do genoma λ .

A disputa entre o repressor λ e Cro é iniciada quando o fago λ infecta uma bactéria normal. A sequência de eventos na disputa é determinada criticamente pela organização dos genes no genoma λ e os promotores e operadores entre os genes *cI* e *cro*. O genoma de λ , com cerca de 50 kb, codifica proteínas que têm papéis na replicação do DNA, recombinação, montagem da partícula de fago e lise celular (Figura 10.26). Essas proteínas são expressas em uma sequência lógica tal que as cópias do genoma são feitas primeiro; essas cópias são então embaladas em partículas virais e, finalmente, a célula hospedeira é lisada para liberação do vírus e começo da infecção de outras células hospedeiras (veja Figura 10.25). A ordem de expressão do gene viral flui do

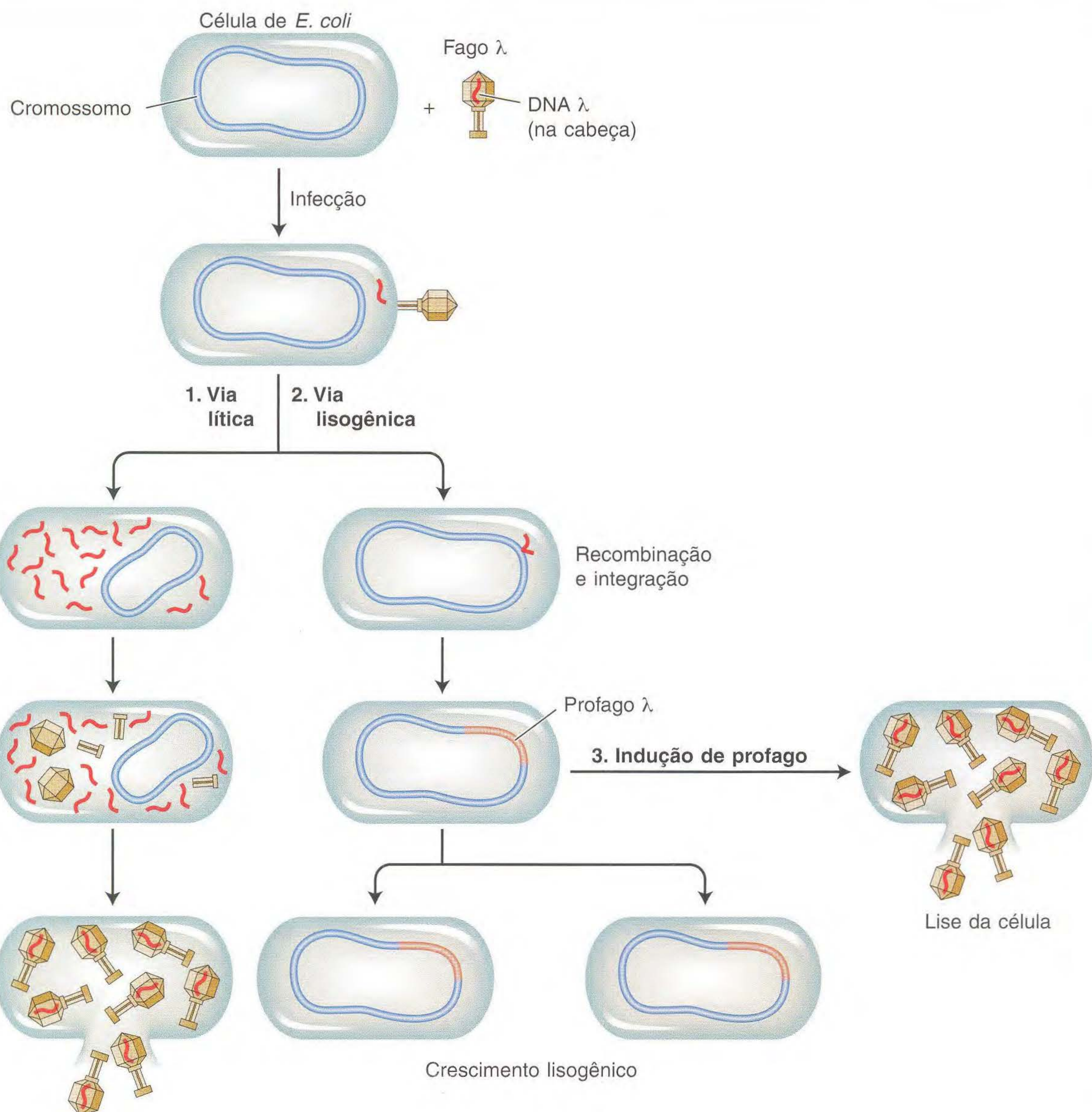
O ciclo vital do bacteriófago λ 

FIG. 10.25 Se o bacteriófago λ entra no ciclo lítico imediatamente ou entra na via lisogênica depende da disponibilidade de recursos. O vírus lisogênico insere seu genoma no cromossomo bacteriano, onde permanece quiescente até que as condições sejam favoráveis.

início da transcrição em dois promotores, P_L e P_R (para promotor à esquerda e direita com relação ao mapa genético). Na infecção, a RNA polimerase inicia a transcrição em ambos os promotores. Olhando o mapa genético (veja Figura 10.26), vemos que, a partir de P_R , *cro* é o primeiro gene transcrito, e, a partir de P_L , *N* é o primeiro gene transcrito. O gene *N* codifica um regulador positivo, mas o mecanismo dessa proteína difere do de outros reguladores que foram considerados até agora. A proteína *N* funciona permitindo que a RNA polimerase continue a transcrever regiões do DNA que de outro modo fariam com que a transcrição termine. Uma proteína reguladora que atua impedindo o término da transcrição é chamada de **antiterminalizador**. Assim, *N* permite a trans-

crição de *cIII* e outros genes à esquerda de *N*, bem como *cII* e outros genes à direita de *cro*. O gene *cII* codifica uma proteína ativadora que se liga a um sítio que promove a transcrição para a esquerda de um promotor diferente, P_{RE} (promotor de estabelecimento do repressor), que ativa a transcrição do gene *cI*. Lembre que o gene *cI* codifica um repressor λ , que impedirá o crescimento lítico.

Antes de a expressão do resto dos genes virais ocorrer, deve ser tomada uma “decisão”: continuar a expressão gênica viral e lisar a célula ou reprimir a via e lisogenizar a célula. A decisão de lisar ou lisogenizar a célula depende da atividade da proteína *cII*. A proteína *cII* é instável, pois é sensível a proteases bacterianas, enzimas que degradam proteínas. As

O ciclo lisogênico *versus* lítico é determinado pela ocupação do repressor nos operadores O_R

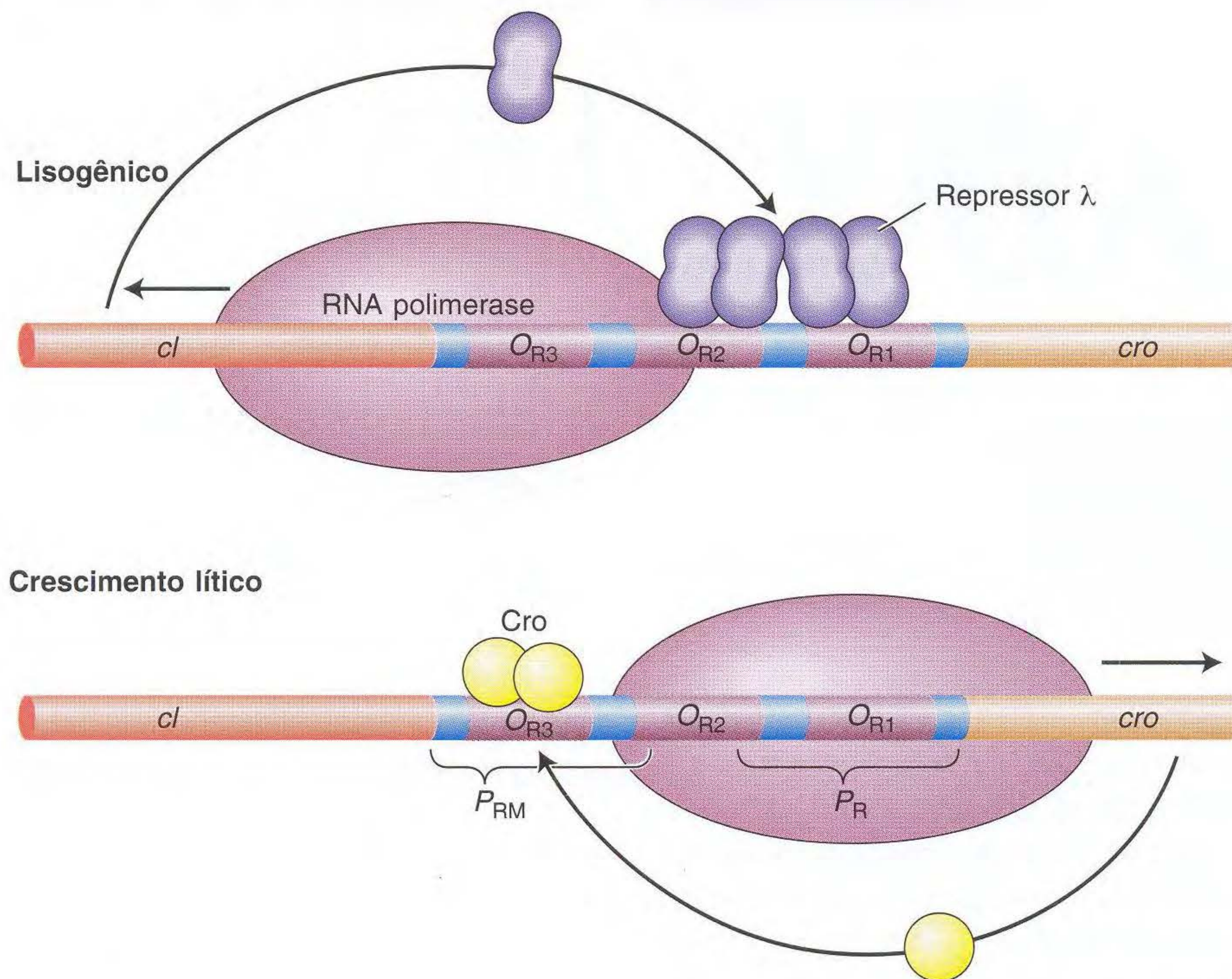


FIG. 10.27 A ligação do repressor λ e Cro aos sítios operadores. A lisogenia é promovida pela ligação do repressor λ a O_{R1} e O_{R2} , que impede a transcrição de P_R . Na indução ou no ciclo lítico, a ligação de Cro a O_{R3} impede a transcrição do gene cI . [De M. Ptashne and A. Gann, *Genes and Signals*, p. 30, Fig. 1-13. © Cold Spring Laboratory Press, 2002.]

P_{RM} (para manutenção do repressor), que dirige a transcrição do gene cI . Os três sítios operadores são similares, mas não idênticos em sequência e, embora Cro e o repressor λ possam cada um ligar-se a um dos operadores, eles fazem isso com afinidades diferentes: o repressor λ liga-se a O_{R1} com a maior afinidade, enquanto Cro liga-se a O_{R3} com a mais alta afinidade. A ocupação do repressor λ de O_{R1} bloqueia a transcrição de P_R e, assim, bloqueia a transcrição dos genes do ciclo lítico. A ocupação de Cro de O_{R3} bloqueia a transcrição de P_{RM} , que bloqueia a manutenção da transcrição de cI e, assim, permite a transcrição dos genes para o ciclo lítico. A ocupação dos sítios operadores determina, portanto, os padrões lítico *versus* lisogênico da expressão do gene λ (Figura 10.27).

Após ter sido estabelecida a lisogenia, ela geralmente é estável. Mas a lisogenia pode ser induzida a entrar no ciclo lítico por várias mudanças ambientais. A luz ultravioleta induz a expressão dos genes hospedeiros. Um dos genes hospedeiros codifica uma proteína, RecA, que estimula a clivagem do repressor λ , prejudicando assim a manutenção da lisogenia e resultando no crescimento lítico. A indução do profago, como Jacob e Monod imaginaram, requer a liberação de um repressor do DNA. O papel fisiológico da luz ultravioleta na indução lisogênica faz sentido nesse tipo de danos por radiação ao DNA hospedeiro e estressa a bactéria; o fago replica-se e deixa a célula estressada, danificada, para outro hospedeiro.

A mudança genética do fago λ ilustra como algumas proteínas reguladoras de ligação ao DNA, agindo por alguns sítios, controlam a expressão de um número muito maior de genes no vírus em um mecanismo "em cascata". Exatamente como os sistemas *lac*, *ara*, *trp* e outros, os estados alternativos da expressão gênica são determinados por sinais fisiológicos.

Ligação específica de sequência de proteínas regulatórias ao DNA

Como um repressor λ e Cro reconhecem operadores diferentes com afinidades diferentes? Essa pergunta dirige nossa atenção para um princípio fundamental no controle do transcrito gênico: a ligação específica de sequência de proteínas regulatórias ao DNA. Para que proteínas individuais se liguem a algumas sequências e não a outras, é necessário especificidade nas interações de cadeias laterais dos aminoácidos da proteína e grupos químicos de bases do DNA. Estudos estruturais detalhados do repressor λ , Cro e outros reguladores bacterianos revelaram como as estruturas tridimensionais dos reguladores e DNA interagem, e como o arranjo de determinados aminoácidos lhes permitem reconhecer sequências específicas de bases.

A análise cristalográfica identificou uma característica estrutural comum dos domínios de ligação ao DNA de λ e Cro. Ambas as proteínas fazem contato com o DNA por meio

Hélice-giro-hélice é um motivo comum de ligação ao DNA

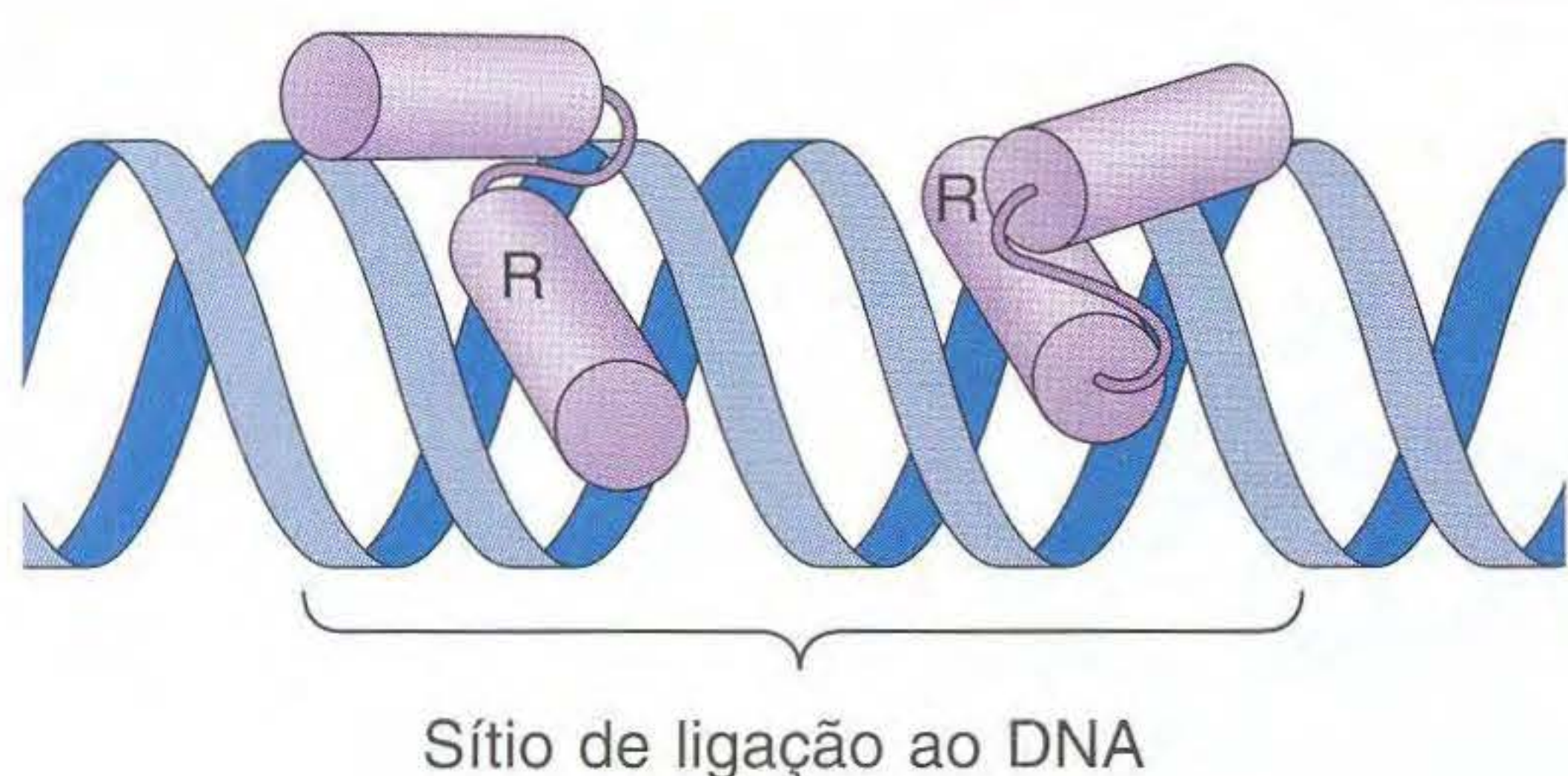


FIG. 10.28 A ligação de um motivo hélice-giro-hélice ao DNA. Muitas proteínas reguladoras ligam-se como dímeros (púrpura) ao DNA. Em cada monômero, a hélice de reconhecimento (R) faz contato com as bases no sulco maior do DNA. [Da Fig. 6-11, p. 494, in J. Watson et al., *Molecular Biology of the Gene*, Fifth Edition. Copyright 2004 by Pearson Education, Inc. Reimpresso com permissão.]

de um domínio *hélice-giro-hélice*, que consiste em duas hélices α ligadas por uma curta região ligadora flexível (Figura 10.28). Uma hélice, a hélice de reconhecimento, ajusta-se ao sulco maior do DNA. Nessa posição, os aminoácidos na face externa da hélice são capazes de interagir com grupos químicos nas bases do DNA. Os aminoácidos específicos na hélice de reconhecimento determinam a afinidade de uma proteína por uma sequência específica de DNA.

As hélices de reconhecimento dos repressores λ e Cro têm estruturas similares e alguns aminoácidos idênticos. As diferenças entre as hélices em aminoácidos-chave determinam suas propriedades de ligação ao DNA. Por exemplo, no repressor λ e proteínas Cro, as cadeias laterais de glutamina e serina fazem contato com as mesmas bases, mas uma alanina, em um repressor λ , e lisina e asparagina, na proteína Cro, impedem diferentes afinidades de ligação para sequências em O_{R1} e O_{R3} (Figura 10.29).

Os repressores Lac e TrpR, bem como o ativador AraC e muitas outras proteínas, também se ligam ao DNA por meio de motivos hélice-giro-hélice de especificidades diferentes, dependendo das sequências de aminoácidos primários de

suas hélices de reconhecimento. Em geral, outros domínios dessas proteínas, tais como os que se ligam a seus respectivos efetores alostéricos, são diferentes.

A especificidade biológica do gene regulador é devida à especificidade química das interações aminoácido-base entre proteínas regulatórias e sequências determinadas do DNA.

► 10.7 Fatores Sigma Alternativos Regulam Grandes Conjuntos de Genes

Até agora, vimos como mudanças únicas podem controlar a expressão de óperons únicos ou dois óperons contendo até duas dúzias de genes. Algumas respostas fisiológicas, tais como a formação de esporos em determinadas bactérias Gram-positivas, requerem a expressão coordenada de um grande grupo de genes não-ligados situados por todo o genoma para causar marcantes mudanças fisiológicas e morfológicas. O processo de esporulação no *Bacillus subtilis* foi analisado em grande detalhe nas últimas décadas. Sob estresse, a bactéria forma esporos que são muito resistentes ao calor e desidratação.

No início do processo de esporulação, a bactéria divide-se assimetricamente, gerando dois componentes de tamanhos desiguais que têm destinos diferentes. O compartimento menor, o pró-esporo, desenvolve-se no esporo. O compartimento maior, a célula-mãe, nutre o esporo em desenvolvimento e sofre lise quando a morfogênese do esporo é completa para liberar o esporo (Figura 10.30a). A dissecação genética desse processo envolveu o isolamento de muitos mutantes que não podem esporular. Investigações detalhadas levaram à caracterização de várias proteínas-chave que regulam diretamente programas de expressão gênica que são específicos do pró-esporo da célula-mãe. Quatro dessas proteínas são fatores σ alternativos.

As cadeias laterais de aminoácidos determinam a especificidade da ligação ao DNA

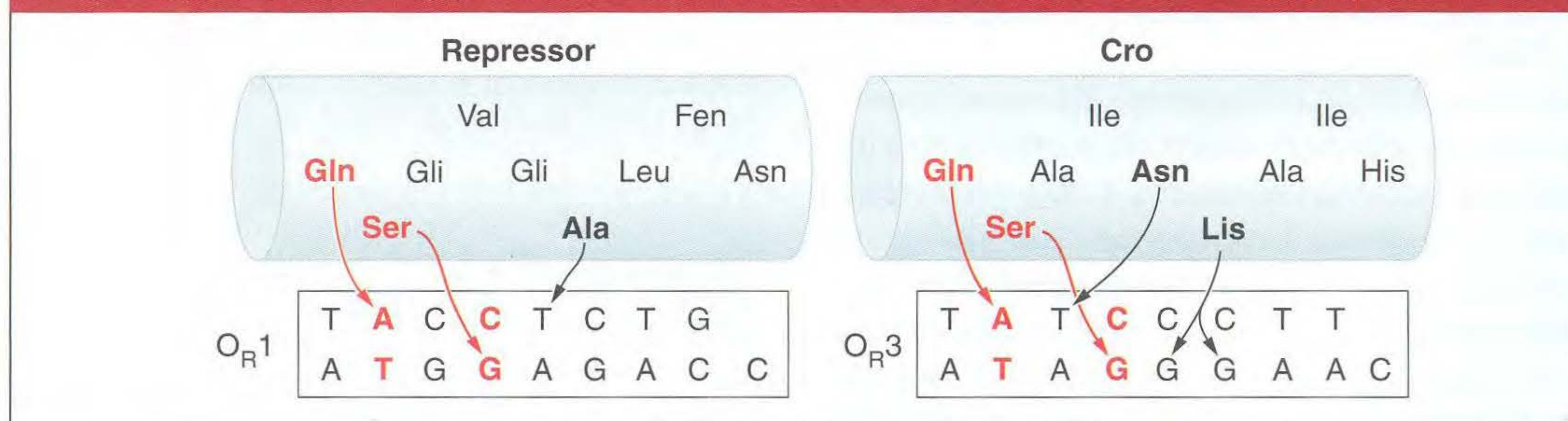


FIG. 10.29 Interações de aminoácidos e bases determinam a especificidade e afinidade das proteínas de ligação ao DNA. As sequências de aminoácidos das hélices de reconhecimento do repressor λ e proteínas Cro são mostradas. As interações de glutamina (Gln), serina (Ser) e alanina (Ala) do repressor λ e as bases do operador O_R determinam a força da ligação. Similarmente, as interações entre glutamina, serina, asparagina (Asn) e lisina (Lis) da proteína Cro medeiam a ligação do operador O_{R3} . Cada sequência de DNA mostrada é a ligada a um monômero individual do respectivo repressor. Ela tem metade do sítio operador ocupado pelo dímero do repressor. [De M. Ptashne, *A Genetic Switch: Phage λ and Higher Organisms*, 2nd ed. Copyright 1992 by Cell Press e Blackwell Scientific.]

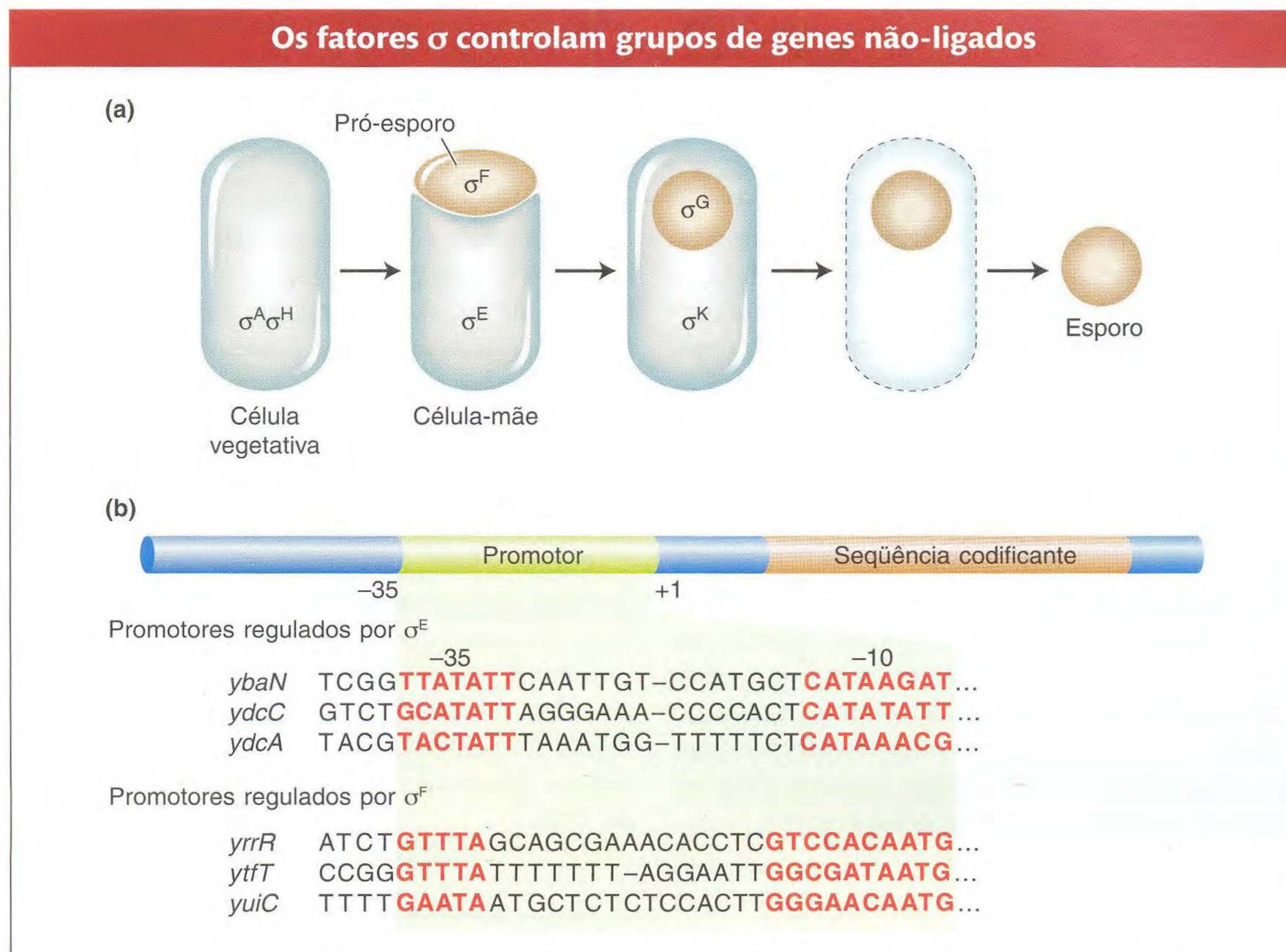


FIG. 10.30 A esporulação em *Bacillus subtilis* é regulada por cascatas de fatores σ . (a) Nas células vegetativas, σ^A e σ^H estão ativos. No início da esporulação, σ^F é ativo no pré-esporo e σ^E é ativo na célula-mãe. Esses fatores σ são então superados por σ^G e σ^K , respectivamente. A célula-mãe lisa e libera o esporo maduro. (b) Os fatores σ^E e σ^F controlam os regulons de muitos genes (*ybaN*, e assim por diante, nesta ilustração). São mostrados três exemplos do grande número de promotores regulados por cada fator σ . Cada fator σ tem uma sequência distinta específica de preferência nas sequências -35 e -10 dos promotores-alvo. [Baseada em dados de P. Eichenberger et al., *J. Mol. Biol.* 327, 2003, 945-972; and S. Wang et al., *J. Mol. Biol.* 358, 2006, 16-37.]

Lembre que o início da transcrição em bactérias inclui a ligação da subunidade σ da RNA polimerase às regiões -35 e -10 dos genes promotores. O fator σ dissocia-se do complexo quando a transcrição começa e é reciclado. Em *B. subtilis*, dois fatores σ , σ^A e σ^H , são ativos nas células vegetativas. Durante a esporulação, um fator σ diferente, σ^F , torna-se ativo no pré-esporo e ativa um grupo de mais de 40 genes. Um gene ativado por σ^F é uma proteína secretada que, por sua vez, dispara o processamento proteolítico do precursor inativo pró- σ^E , um fator σ distinto na célula-mãe. O fator σ^E é necessário para ativar conjuntos de genes na célula-mãe. Dois fatores σ adicionais, σ^K e σ^G , são subsequentemente ativados na célula-mãe e pré-esporo, respectivamente (Figura 10.30a). A expressão de fatores σ distintos permite a transcrição coordenada de conjuntos diferentes de genes, ou **regulons**, por uma única RNA polimerase.

Como esses fatores σ alternativos controlam aspectos diferentes do processo de esporulação? As respostas ficaram claras com o advento de novos enfoques para caracterização da expressão de todos os genes em um genoma (veja Seção 13.7). Hoje é possível monitorar a transcrição de cada gene de *B. subtilis* durante o crescimento vegetativo e a formação de esporos, e em compartimentos diferentes do esporo. Várias centenas de genes foram identificados desse modo,

sendo ativados transcricionalmente ou reprimidos durante a formação de esporos.

Como os diferentes conjuntos de genes são controlados por cada fator σ ? Cada fator σ tem propriedades de ligação ao DNA específicas de sequências. Os óperons ou genes individuais regulados por fatores σ específicos têm sequências características nas regiões -35 e -10 de seus promotores que são ligadas por um fator σ e não outros (Figura 10.30b). Por exemplo, σ^E liga-se a, pelo menos, 121 promotores, dentro de 34 óperons e 87 genes individuais, para regular mais de 250 genes, e σ^F liga-se a, pelo menos, 36 promotores para regular 48 genes.

A expressão sequencial de fatores σ alternativos que reconhecem sequências promotoras alternativas fornece a expressão coordenada de um grande número de óperons independentes e genes não-ligados durante o programa desenvolvimental da esporulação.

Fatores σ alternativos também têm papéis importantes na virulência de patógenos humanos. Por exemplo, as bactérias do gênero *Clostridium* produzem toxinas potentes que são responsáveis por doenças graves, tais como botulismo, tétano e gangrena. Genes principais de toxinas de *C. botulinum*,

C. tetani e *C. perfringens* recentemente foram descobertos, sendo controlados por fatores σ alternativos, correlatos, que reconhecem seqüências similares nas regiões -35 e -10 dos

genes de toxina. A compreensão dos mecanismos de regulação dos genes de toxina pode levar a um novo modo de prevenção de doenças e terapia.

► Resumo

A regulação gênica geralmente é mediada por proteínas que reagem a sinais ambientais aumentando ou diminuindo as taxas de transcrição de genes específicos. A lógica dessa regulação é direta. Para que a regulação opere apropriadamente, as proteínas regulatórias têm sensores internos que, continuamente, monitoram as condições ambientais. As atividades dessas proteínas dependeriam, então, das condições ambientais apropriadas.

Nas bactérias e seus vírus, o controle de vários genes estruturais pode ser coordenado aglomerando genes em óperons no cromossomo, de modo que eles são transcritos em mRNA multigênicos. O controle coordenado simplifica a tarefa das bactérias porque um grupo de sítios regulatórios por óperon é suficiente para regular a expressão de todos os genes do óperon. Alternativamente, o controle coordenado também pode ser obtido por distintos fatores σ que regulam simultaneamente dúzias de promotores independentemente.

No controle regulador negativo, uma proteína repressora bloqueia a transcrição ligando-se ao DNA no sítio operador. O controle regulador negativo é exemplificado pelo sistema *lac*. A regulação negativa é um modo bem direto de o sistema *lac* desligar genes na ausência de açúcares apropriados no ambiente. No controle regulador positivo, são necessários fatores proteicos para ativar a transcrição. Alguns controles gênicos procarióticos, tais como o da repressão catabólica, operam por controle genético positivo.

Muitas proteínas regulatórias são membros de famílias de proteínas que têm motivos de ligação ao DNA muito similares, tais como o domínio hélice-giro-hélice. Outras partes das proteínas, tais como seus domínios de interação proteína-proteína, tendem a ser menos similares. A especificidade da regulação gênica depende das interações químicas entre as cadeias laterais de aminoácidos e grupos químicos nas bases do DNA.

► Termos-chave

ação cis	diplóide parcial	promotor
ação trans	domínio de ligação ao DNA	proteína ativadora do catabolismo (CAP)
adenosina monofosfato cíclica (cAMP)	efetor alostérico	regulon
antiterminalizador	genes controlados coordenadamente	repressão catabólica
atenuação	indução	repressor
atenuador	indutor	seqüência líder
ativador	iniciador	sítio alostérico
ciclo lisogênico	mudança genética	transição alostérica
ciclo lítico	mutação constitutiva	
controle negativo	operador	
controle positivo	óperon	

► Problemas Resolvidos

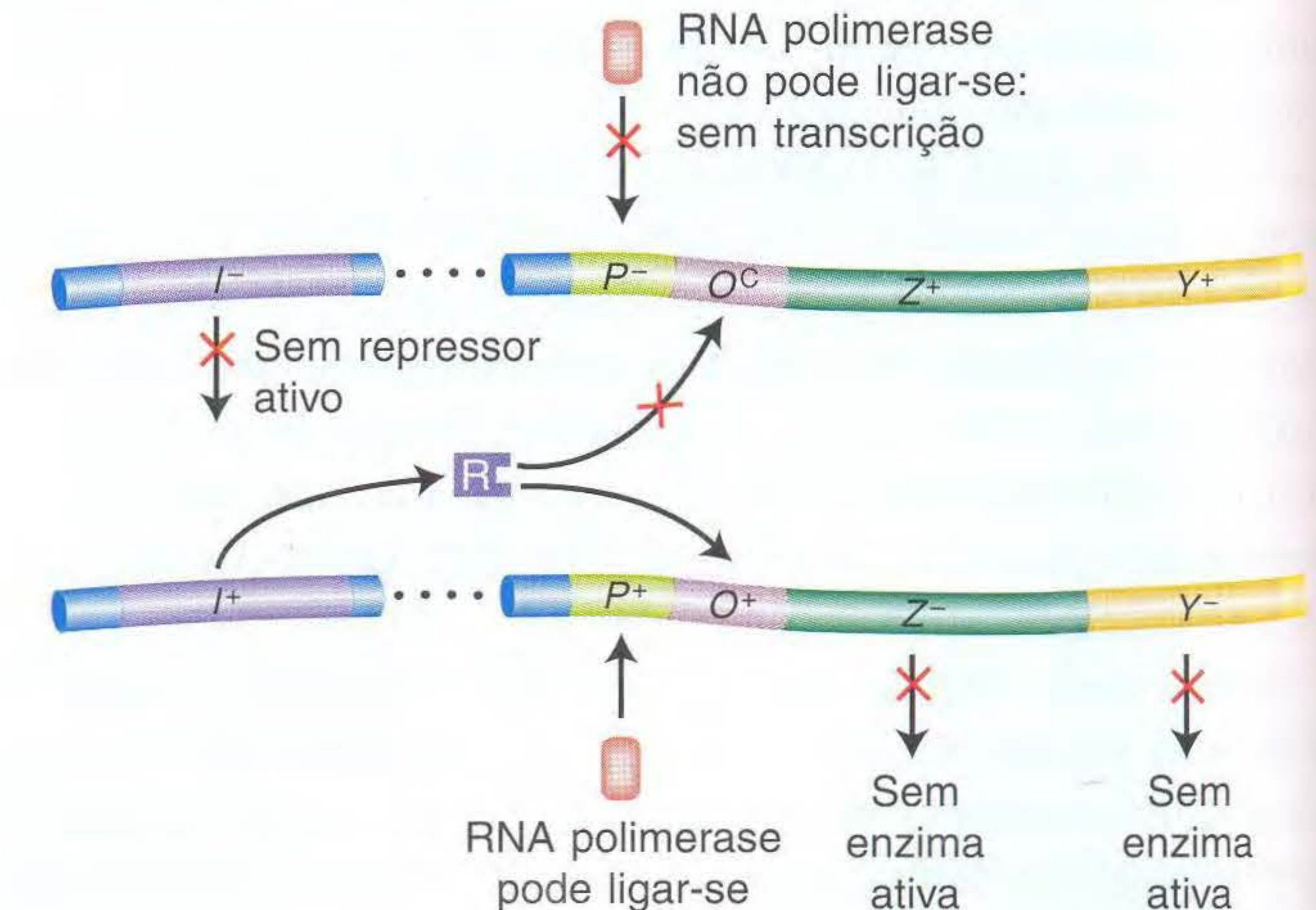
Este conjunto de quatro problemas resolvidos, que é similar ao Problema 4 dos Problemas Básicos ao final deste capítulo, é destinado a testar a compreensão do modelo do óperon. Aqui, damos vários diplóides e pedimos determinar se os produtos dos genes Z e Y são feitos na presença ou ausência de um indutor. Use um quadro similar ao do Problema 4 como base para suas respostas, mas os títulos das colunas serão os seguintes:

Genótipo	Gene Z		Gene Y	
	Sem indutor	Indutor	Sem indutor	Indutor

Problema resolvido 1. $I^- P^- O^C Z^+ Y^+$
 $I^+ P^+ O^+ Z^- Y^-$

SOLUÇÃO

Um modo para abordar esses problemas é primeiro considerar cada cromossomo separadamente e, então, construir um diagrama. A ilustração seguinte diagrama esse diplóide:



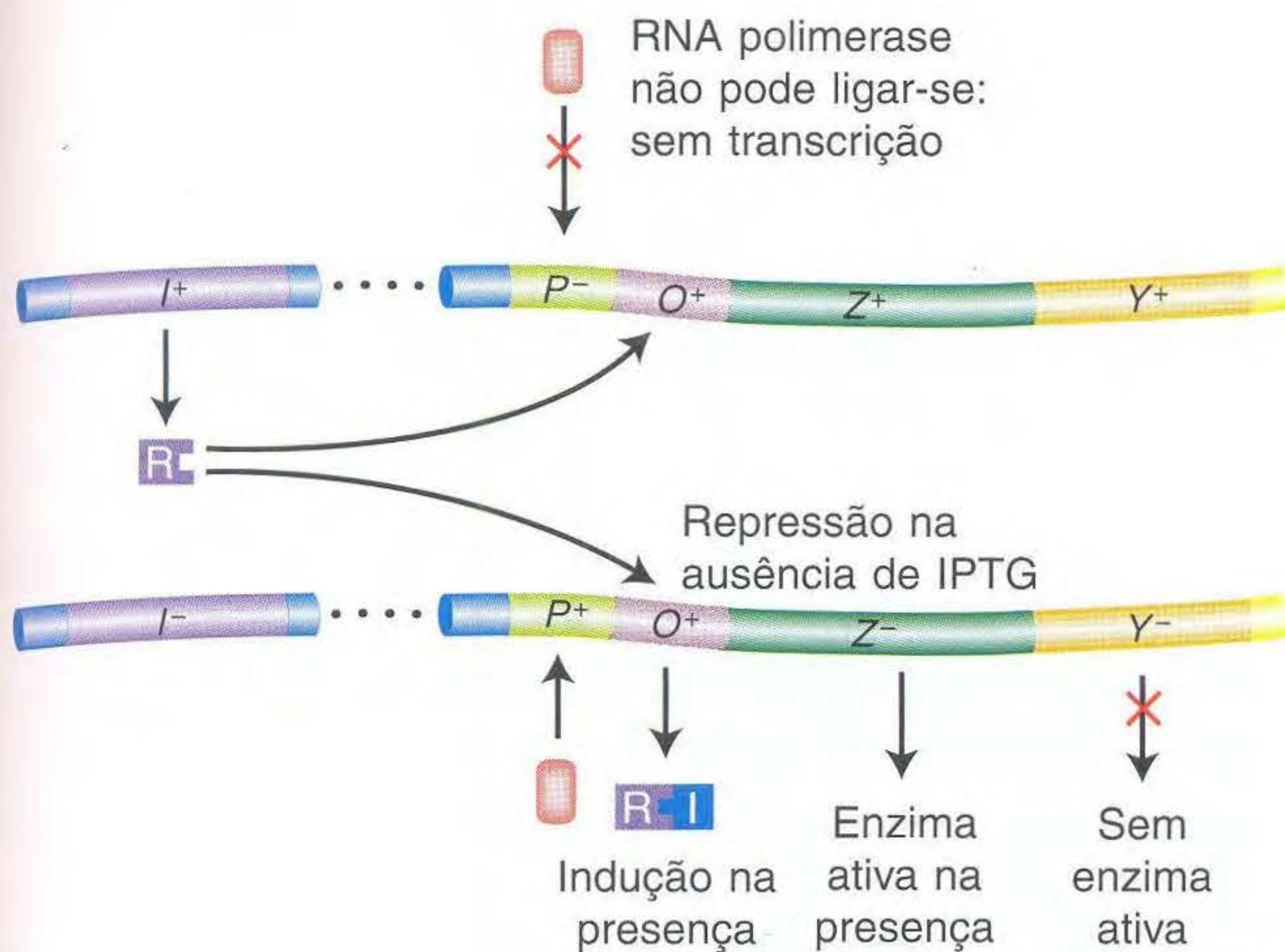
O primeiro cromossomo é P^- , e, assim, o transcrito está bloqueado e nenhuma enzima Lac pode ser sintetizada a partir dele. O segundo cromossomo (P^+) pode ser transcrito, e,

assim, a transcrição é repressível (O^+). Entretanto, os genes estruturais ligados ao bom promotor são defectivos; logo, nenhum produto ativo Z ou Y pode ser gerado. Os símbolos a serem adicionados a seu quadro são “-, -, -, -”.

Problema resolvido 2. $\frac{I^+ P^- O^+ Z^+ Y^+}{I^- P^+ O^+ Z^+ Y^-}$

SOLUÇÃO

O primeiro cromossomo é P^- , e, assim, nenhuma enzima pode ser produzida por ele. O segundo cromossomo é O^+ , e, assim, a transcrição é reprimida pelo repressor fornecido pelo primeiro cromossomo, que pode agir em trans através do citoplasma. Entretanto, apenas o gene Z desse cromossomo está intacto. Portanto, na ausência de um indutor, não é feita nenhuma enzima; na presença de um indutor, apenas o produto gênico Z, a β galactosidase, é gerado. Os símbolos para adicionar ao quadro são “-, +, -, -”.



Problema resolvido 3. $\frac{I^+ P^+ O^C Z^- Y^+}{I^+ P^- O^+ Z^+ Y^-}$

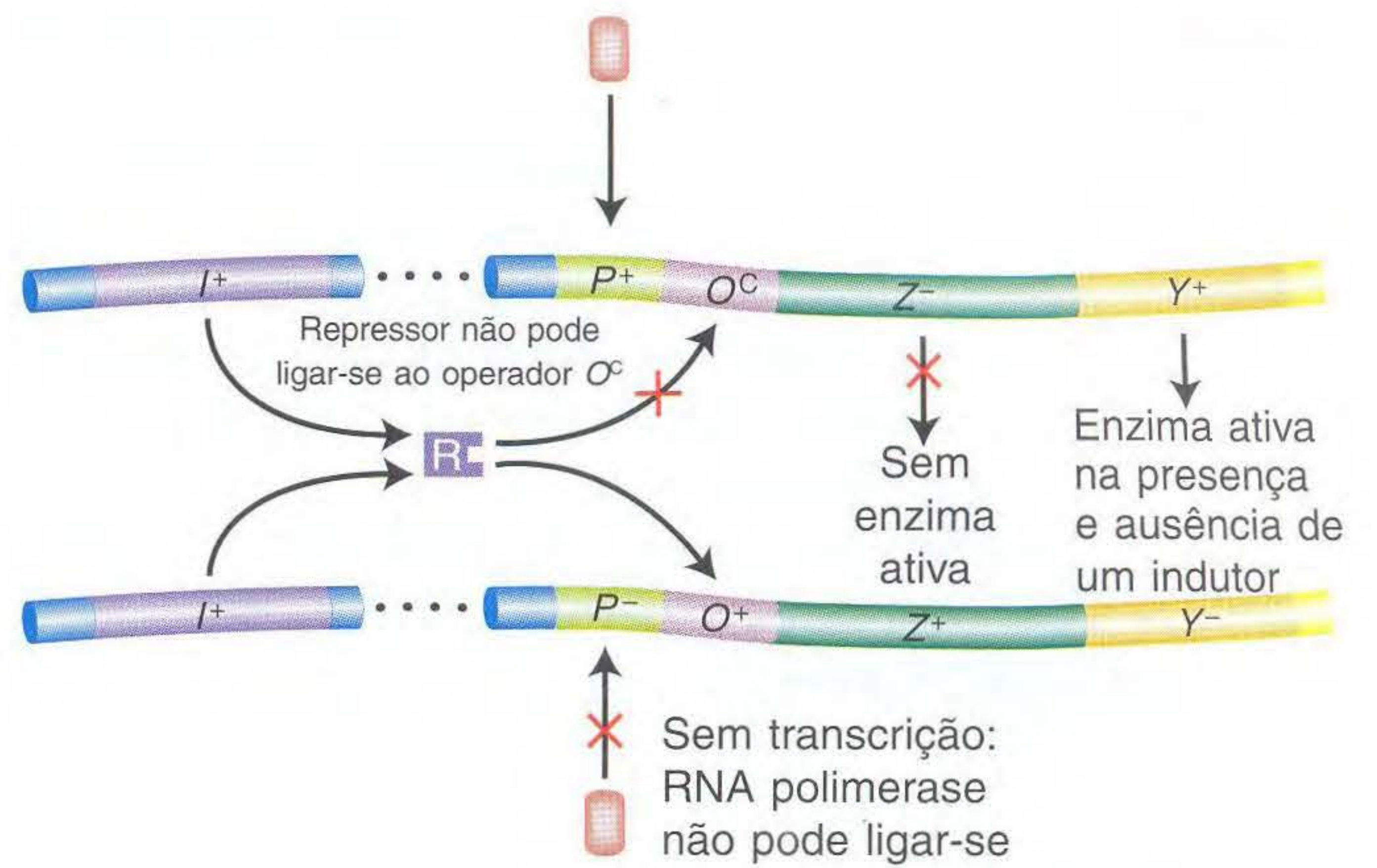
SOLUÇÃO

Como o segundo cromossomo é P^- , precisamos considerar apenas o primeiro cromossomo. Esse cromossomo é O^C , e, assim, a enzima é feita na ausência de um indutor, embora, devido à mutação Z^- , apenas a permease ativa é gerada. As entradas no quadro devem ser “-, -, +, +”.

Problemas

PROBLEMAS BÁSICOS

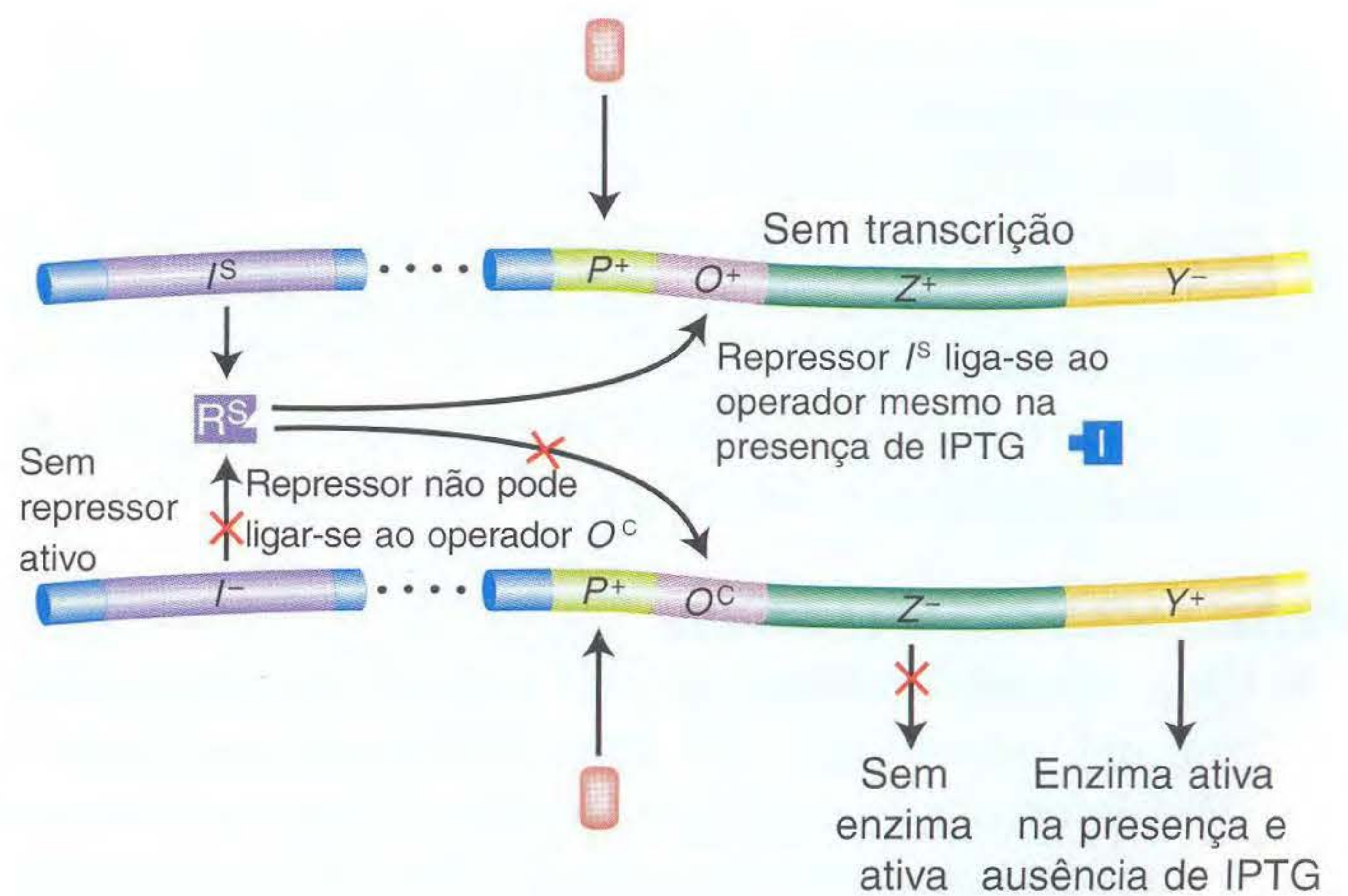
1. Explique por que as mutações I^- no sistema *lac* são normalmente recessivas às mutações I^+ e por que as mutações I^+ são recessivas às mutações I^S .
2. O que queremos dizer quando falamos que as mutações O^C no sistema *lac* são de ação cis?
3. Os genes mostrados no quadro seguinte são do sistema do óperon *lac* de *E. coli*. Os símbolos *a*, *b* e *c* representam o gene repressor (*I*), o operador (*O*) e o gene estrutural (*Z*) para β -galactosidase, embora não necessariamente nessa ordem. Além disso, a ordem na qual os símbo-



Problema resolvido 4. $\frac{I^S P^+ O^+ Z^+ Y^-}{I^- P^+ O^C Z^- Y^+}$

SOLUÇÃO

Na presença do repressor I^S , todos os operadores tipo selvagem são desligados, tanto com e quanto sem um indutor. Portanto, o primeiro cromossomo é incapaz de produzir uma enzima. Entretanto, o segundo cromossomo tem um operador alterado (O^C) e pode produzir enzima tanto na ausência quanto na presença de um indutor. Apenas o gene Y é tipo selvagem no cromossomo O^C , e, assim, só a permease é produzida constitutivamente. As entradas no quadro devem ser “-, -, +, +”.



los são escritos nos genótipos não é necessariamente a sequência real no óperon *lac*.

Atividade (+) ou inatividade (-) do gene Z

Genótipo	Indutor ausente	Indutor presente
$a^- b^+ c^+$	+	+
$a^+ b^+ c^-$	+	+
$a^+ b^- c^-$	-	-
$a^+ b^- c^+/a^- b^+ c^-$	+	+
$a^+ b^+ c^+/a^- b^- c^-$	-	+
$a^+ b^+ c^-/a^- b^- c^+$	-	+
$a^- b^+ c^+/a^+ b^- c^-$	+	+

- a. Diga qual símbolo (*a*, *b* ou *c*) representa cada um dos genes *lac*, *I*, *O* e *Z*.
- b. No quadro, um expoente menos em um símbolo de um gene indica apenas um mutante, mas você sabe que alguns comportamentos mutantes nesse sistema recebem designações mutantes especiais. Use os símbolos gênicos convencionais para o óperon *lac* para designar cada genótipo no quadro.
- (O problema 3 é de J. Kuspira and G.W. Walker, *Genetics: Questions and Problems*. Copyright 1973 by McGraw-Hill.)
4. O mapa do óperon *lac* é

POZY

Genótipo	β-galactosidase		Permease	
	Sem lactose	Lactose	Sem lactose	Lactose
<i>I</i> ⁺ <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> ⁺ <i>Y</i> ⁺ / <i>I</i> ⁺ <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> ⁺ <i>Y</i> ⁺	–	+	–	+
a. <i>I</i> [–] <i>P</i> ⁺ <i>O</i> ^c <i>Z</i> ⁺ <i>Y</i> [–] / <i>I</i> ⁺ <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> [–] <i>Y</i> ⁺				
b. <i>I</i> ⁺ <i>P</i> [–] <i>O</i> ^c <i>Z</i> [–] <i>Y</i> ⁺ / <i>I</i> [–] <i>P</i> ⁺ <i>O</i> ^c <i>Z</i> ⁺ <i>Y</i> [–]				
c. <i>I</i> ^s <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> ⁺ <i>Y</i> [–] / <i>I</i> ⁺ <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> [–] <i>Y</i> ⁺				
d. <i>I</i> ^s <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> ⁺ <i>Y</i> ⁺ / <i>I</i> [–] <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> ⁺ <i>Y</i> ⁺				
e. <i>I</i> [–] <i>P</i> ⁺ <i>O</i> ^c <i>Z</i> ⁺ <i>Y</i> [–] / <i>I</i> [–] <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> [–] <i>Y</i> ⁺				
f. <i>I</i> [–] <i>P</i> [–] <i>O</i> ⁺ <i>Z</i> ⁺ <i>Y</i> ⁺ / <i>I</i> [–] <i>P</i> ⁺ <i>O</i> ^c <i>Z</i> ⁺ <i>Y</i> [–]				
g. <i>I</i> ⁺ <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> [–] <i>Y</i> ⁺ / <i>I</i> [–] <i>P</i> ⁺ <i>O</i> ⁺ <i>Z</i> ⁺ <i>Y</i> [–]				

5. Explique as diferenças fundamentais entre o controle negativo e o controle positivo.
6. Os mutantes que são *lacY*[–] conservam a capacidade de sintetizar β-galactosidase. Entretanto, muito embora o gene *lacI* ainda esteja intacto, a β-galactosidase não pode mais ser induzida pela adição de lactose ao meio. Explique.
7. Quais são as analogias entre os mecanismos que controlam o óperon *lac* e os que controlam as mudanças genéticas do fago λ?
8. Compare o arranjo dos sítios de ação cis nas regiões de controle do óperon *lac* e fago λ.

PROBLEMAS DESAFIADORES

9. Uma mutação curiosa em *lacI* resulta nos repressores com um aumento de 100 vezes de ligação tanto com o DNA operador quanto não-operador. Esses repressores apresentam uma curva de indução “reversa”, permitindo a síntese de β-galactosidase na ausência de um indutor (IPTG), mas reprimindo parcialmente a expressão de β-galactosidase na presença de IPTG. Como você pode explicar isso? (Note que, quando IPTG se liga a um repressor, ele não destrói completamente a afinidade pelo operador, mas reduz em 1.000 vezes a afinidade. Além disso, à medida que as células se dividem e são gerados novos operadores pela síntese de filamentos-filhos, o repressor deve encontrar os novos operadores procurando ao longo do DNA, ligando-se rapidamente a seqüências não-operadoras e dissociando-se delas.)
10. Em *Neurospora*, todos os mutantes que afetam as enzimas carbamil fosfato sintetase e aspartato transcarbamilase estão mapeados no locus de *pyr-3*. Se você induz

A região promotora (*P*) é o ponto de início da transcrição, através da ligação da molécula de RNA polimerase antes da produção real de mRNA. Promotores mutacionalmente alterados (*P*[–]) não podem ligar a molécula de RNA polimerase. Algumas previsões podem ser feitas sobre o efeito das mutações *P*[–]. Use suas previsões e seus conhecimentos do sistema lactose para completar o quadro seguinte. Insira um “+” onde uma enzima é produzida, e um “–” onde nenhuma enzima é produzida. O primeiro foi feito como exemplo.

- mutações em *pyr-3* por ICR-170 (um mutágeno químico), você descobre que ou ambas as funções enzimáticas estão faltando ou apenas a função transcarbamilase está faltando; em nenhum caso a atividade de sintetase está faltando quando a atividade de transcarbamilase está presente. (ICR-170 supostamente induz mudanças de matriz de leitura.) Interprete esses resultados com relação a um possível óperon.
11. Algumas mutações em *lacI* eliminam a ligação do operador pelo repressor Lac, mas não afetam a agregação de subunidades para fazer um tetrâmero, a forma ativa do repressor. Esses mutantes são parcialmente dominantes em relação ao tipo selvagem. Você pode explicar o fenótipo parcialmente dominante *I*[–] nos heterodiplóides *I*[–]/*I*⁺?
12. Você está examinando a regulação do óperon de lactose na bactérias *Escherichia coli*. Você isola sete novas linhagens mutantes independentes que não têm os produtos proteicos de todos os três genes estruturais. Você suspeita que algumas dessas mutações são *lacI*^s e outras mutações são alterações que impedem a ligação da RNA polimerase à região do promotor. Usando qualquer genótipo haplóide e parcialmente diplóide que você imagine ser necessário, descreva um conjunto de genótipos que lhe permitirão distinguir as classes *lacI* e *lacP* de mutações não-indutíveis.
13. Você está estudando as propriedades de um novo tipo de mutação regulatória do óperon de lactose. Essa mutação, chamada *S*, leva a uma completa repressão dos genes *lacZ*, *lacY* e *lacA*, independentemente da presença de um indutor (lactose). Os resultados dos estudos dessa muta-

- ção em diplóides parciais demonstram que essa mutação é totalmente dominante em relação ao tipo selvagem. Quando você trata bactérias da linhagem mutante *S* com um mutágeno e seleciona bactérias mutantes que podem expressar as enzimas codificadas pelos genes *lacZ*, *lacY* e *lacA* na presença de lactose, algumas das mutações mapeiam na região do operador *lac* e outras no gene repressor *lac*. Com base em seus conhecimentos do óperon de lactose, forneça uma explicação genética molecular para todas essas propriedades da mutação *S*. Inclua uma explicação da natureza constitutiva das “mutações reversas”.
14. O óperon *trp* em *E. coli* codifica enzimas essenciais para a biossíntese de triptofano. O mecanismo geral para o controle do óperon *trp* é similar ao observado com o óperon *lac*: quando o repressor se liga ao operador, a transcrição é evitada; quando o repressor não se liga ao operador, a transcrição continua. A regulação do óperon *trp* difere da regulação do óperon *lac* do seguinte modo: as enzimas codificadas pelo óperon *trp* não são sintetizadas quando o triptofano está presente, mas sim quando este está ausente. No óperon *trp*, o repressor tem dois sítios de ligação: um para o DNA e outro para a molécula efetora, o triptofano. O repressor *trp* deve primeiro ligar-se a uma molécula de triptofano antes que possa ligar-se efetivamente ao operador *trp*.
- a. Desenhe um mapa do óperon de triptofano, indicando o promotor (*P*), o operador (*O*) e o primeiro gene estrutural do óperon de triptofano (*trpA*). Em seu desenho, indique onde, no DNA, a proteína repressora se liga quando está ligada ao triptofano.
- b. O gene *trpR* codifica o repressor; *trpO* é o operador; *trpA* codifica a enzima triptofano sintetase. Um repressor *trpR*² não pode ligar-se ao triptofano; um operador *trpO*² não pode ligar-se ao repressor; e a enzima codificada pelo gene mutante *trpA*² é completamente inativa.

va. Você espera encontrar triptofano sintetase em cada uma das seguintes linhagens mutantes quando as células estão crescendo na presença de triptofano? Em sua ausência?

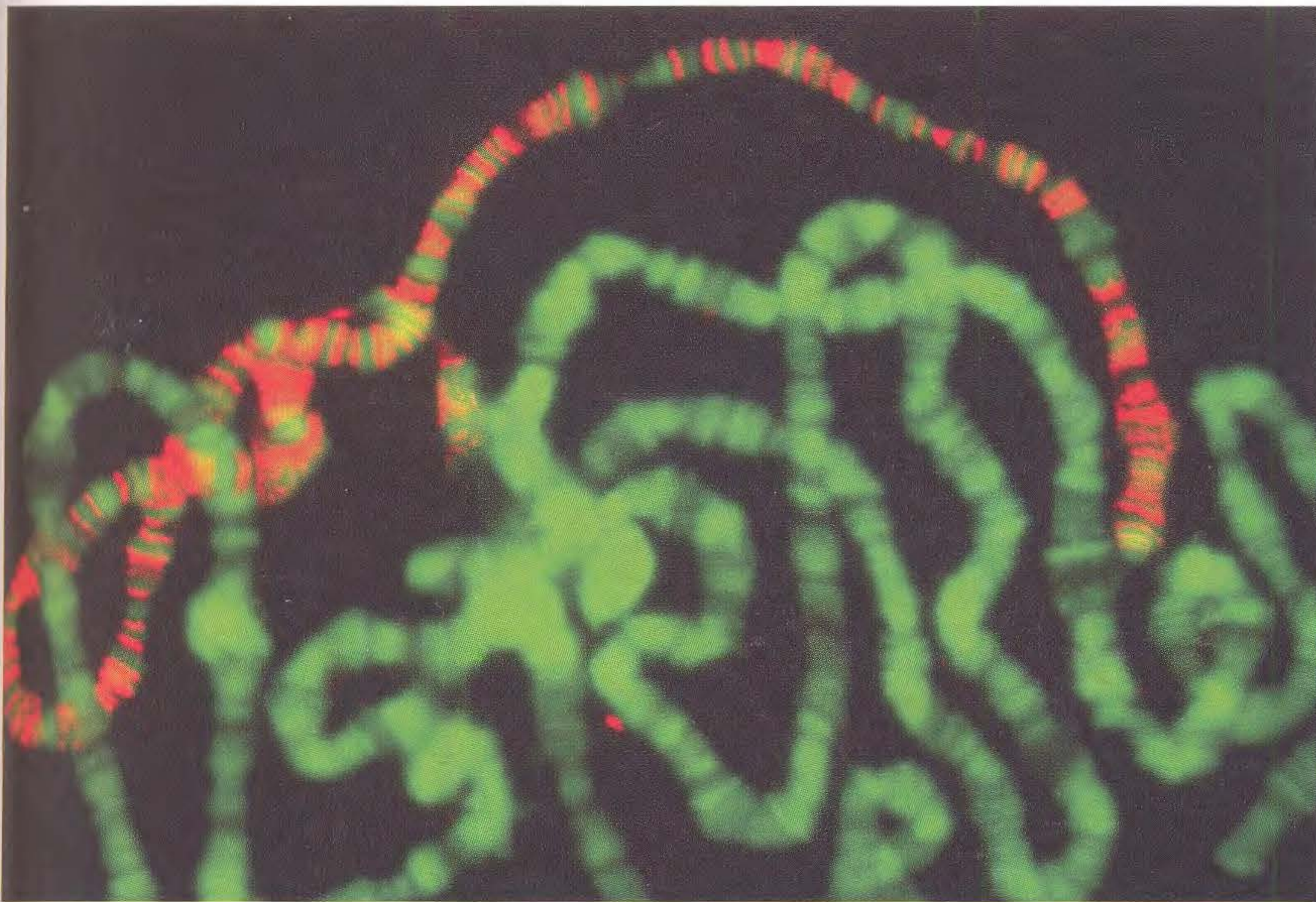
- i. *R*⁺ *O*⁺ *A*⁺ (tipo selvagem)
ii. *R*⁻ *O*⁺ *A*⁺/*R*⁺ *O*⁺ *A*⁻
iii. *R*⁺ *O*⁻ *A*⁺/*R*⁺ *O*⁺ *A*⁻
15. A atividade da enzima β-galactosidase produzida pelas células tipo selvagem cultivadas em meio suplementado com diferentes fontes de carbono é medida. Em unidades relativas, foram encontrados os seguintes níveis de atividade:

Glicose	Lactose	Lactose + glicose
0	100	1

Preveja os níveis relativos de atividade de β-galactosidase nas células cultivadas sob condições similares quando as células são *lacI*⁻, *lacI*^s, *lacO*^c e *crp*⁻.

16. Um bacteriófago λ é visto sendo capaz de lisogenizar sua *E. coli* hospedeira a 30°C, mas não a 42°C. Que genes podem ser mutantes nesse fago?
17. O que aconteceria com a habilidade do bacteriófago λ em lisar uma célula hospedeira se tivesse uma mutação no sítio de ligação *O*_R para a proteína Cro? Por quê?
18. Contraste os efeitos das mutações nos genes codificantes de fatores σ específicos da esporulação com mutações nas regiões -35 e -10 dos promotores dos genes em seus regulons.
- a. As mutações funcionais nos genes de fator σ ou nos promotores individuais têm o maior efeito na esporulação?
- b. Com base nas seqüências mostradas na Figura 10.30b, você esperaria que todas as mutações de ponto em -35 ou -10 afetem a expressão gênica?

Regulação da Expressão Gênica em Eucariontes



O complexo MSL acentua a expressão gênica no cromossomo X. O complexo MSL (indicado pela cor laranja) liga-se apenas ao cromossomo X nos machos de *Drosophila*. Essa imagem é uma coloração com imunofluorescência indireta de uma dispersão cromossômica de uma glândula salivar de uma larva macho ao anti-soro MSL1. [De J. Lucchesi, W. Kelly, and B. Panning. "Chromatin Remodeling in Dosage Compensation," *Annu. Rev. Genet.* 39, 2005, 615–651.]

A clonagem de Dolly, uma ovelha, foi relatada em todo o mundo em 1996. Dolly desenvolveu-se de núcleos somáticos adultos que tinham sido implantados em ovócitos enucleados (sem núcleos). Mais recentemente, vacas, porcos, camundongos e outros mamíferos também foram clonados com o uso de tecnologia similar (Figura 11.1). A clonagem bem-sucedida de Dolly foi uma grande surpresa para a comunidade científica porque a clonagem de mamíferos a partir de células somáticas era tida como impossível. Um motivo para o ceticismo inicial foi que a formação de gametas masculinos e femininos (espermatozoides e ovócitos) era conhecida como incluindo modificações específicas do sexo nos respectivos genomas que resultaram em padrões específicos do sexo de expressão gênica. Como tal, Dolly é simbólica para demonstrar o quanto progredimos na compreensão de aspectos da regulação gênica eucariótica, como o controle global da expressão gênica exemplificado pelo desenvolvimento de gametas. Entretanto, para cada clone bem-sucedido, inclusive Dolly, existem muitos mais, talvez centenas de embriões que não se desenvolvem em prole viável. A taxa extremamente alta de falhas destaca o quanto ainda falta para ser decifrado sobre a regulação gênica eucariótica.

Neste capítulo, examinaremos a regulação gênica nos eucariontes. De muitos modos, nossa visão da regulação gênica será um estudo de contrastes. Nas

11

Perguntas Fundamentais

- Quais são os mecanismos moleculares de regulação gênica nos eucariontes?
- Como os eucariontes geram muitos padrões diferentes de expressão gênica com um número limitado de proteínas reguladoras?
- Que papel a cromatina tem na regulação gênica eucariótica?
- Quais são as marcas epigenéticas e como elas influenciam a expressão gênica?

Tópicos

- 11.1 Regulação transcricional em eucariontes: visão geral
- 11.2 Lições de leveduras: o sistema GAL
- 11.3 Cromatina dinâmica e regulação gênica eucariótica
- 11.4 Mecanismo de ação acentuadora
- 11.5 Imprinting genômico
- 11.6 Domínios de cromatina e sua herança

O primeiro mamífero clonado



FIG. 11.1 O primeiro mamífero clonado foi uma ovelha chamada Dolly. [PHOTOTAKE Inc./Alamy.]

bactérias, você aprendeu como as atividades das mudanças genéticas em geral eram governadas por proteínas ativadoras ou repressoras únicas, e como o controle de grupos de genes foi obtido por sua organização em óperons ou pela atividade de fatores σ específicos (veja o Capítulo 10). As expectativas iniciais eram de que a expressão gênica eucariótica seria regulada por meios similares. Nos eucariontes, entretanto, a maioria dos genes não são encontrados em óperons. Além disso, veremos que as proteínas e as seqüências de DNA que participam da regulação gênica eucariótica são mais numerosas. Em geral, muitas proteínas de ligação ao DNA atuam em uma única mudança, com muitos “interruptores” separados por gene, e as seqüências regulatórias desses interruptores quase sempre estão situadas longe dos promotores. Uma importante diferença adicional entre bactérias e eucariontes é que o acesso aos promotores gênicos eucarióticos é restrito pela cromatina. A regulação gênica em eucariontes requer a atividade de grandes complexos de proteína que promovem ou restringem o acesso a promotores gênicos pela RNA polimerase. Este capítulo dará as bases essenciais para a compreensão da regulação espaço-temporal da expressão gênica que coreografa o processo de desenvolvimento descrito no Capítulo 12.

► 11.1 Regulação Transcricional em Eucariontes: Visão Geral

As propriedades biológicas de cada tipo de célula eucariótica são amplamente determinadas pelas proteínas expressas dentro dela. Essa constelação de proteínas expressas determina muito da arquitetura celular, suas atividades enzimáticas, suas interações com seu ambiente, além de muitas outras propriedades fisiológicas. Entretanto, em um momento qualquer da história de vida de uma célula, apenas uma fração

dos RNA e proteínas codificadas em seu genoma é expressa. Em momentos diferentes, o perfil dos produtos gênicos expressos pode diferir marcadamente, tanto com relação às proteínas que são expressas como em que níveis. De que modo são gerados esses perfis específicos?

Como se poderia esperar, se o produto final é uma proteína, a regulação poderia ser obtida controlando a transcrição do DNA em RNA, ou a tradução do RNA em proteína. De fato, a regulação gênica ocorre em muitos níveis, inclusive ao nível de mRNA (por alterações na recomposição ou estabilidade do mRNA) e após a tradução (por modificação das proteínas). Entretanto, a maioria da regulação é tida como ocorrendo ao nível da transcrição gênica; assim, neste capítulo, o foco primário é na regulação da transcrição. O mecanismo básico em funcionamento é que os sinais moleculares de fora ou de dentro da célula levam à ligação de proteínas reguladoras a sítios específicos do DNA fora das regiões codificantes de proteína, e a ligação dessas proteínas modula a taxa de transcrição. Essas proteínas podem, direta ou indiretamente, ajudar a RNA polimerase na ligação a seu sítio de início da transcrição, o *promotor*, ou podem reprimir a transcrição evitando a ligação da RNA polimerase.

Embora as bactérias e eucariontes tenham muito da lógica da regulação gênica em comum, existem algumas diferenças fundamentais nos mecanismos subjacentes e maquinaria. Ambos usam proteínas de ligação ao DNA com seqüências específicas para modular o nível de transcrição. Entretanto, os genomas eucarióticos são maiores e sua gama de propriedades é maior do que a de bactérias. Inevitavelmente, sua regulação é mais complexa, requerendo mais tipos de proteínas reguladoras e mais tipos de interações com regiões reguladoras adjacentes no DNA. A diferença mais importante é que o DNA eucariótico é compactado em *nucleossomos*, formando *cromatina*, enquanto o DNA bacteriano não tem nucleossomos. Nos eucariontes, a estrutura da cromatina é dinâmica e é um ingrediente essencial na regulação gênica.

Em geral, o estado básico de um gene bacteriano é “ligado”. Assim, a RNA polimerase geralmente pode se ligar a um promotor quando nenhuma outra proteína reguladora está por perto para se ligar ao DNA. Nas bactérias, o início da transcrição é evitado ou reduzido se a ligação da RNA polimerase estiver bloqueada, geralmente pela ligação de uma proteína reguladora repressora. As proteínas reguladoras ativadoras aumentam a ligação da RNA polimerase a promotores onde pouca ajuda é necessária. Em contraste, o estado básico em eucariontes é “desligado”. Portanto, a maquinaria transcricional (incluindo a RNA polimerase II e fatores gerais de transcrição associados) não podem ligar-se ao promotor na ausência de outras proteínas reguladoras (Figura 11.2). Em muitos casos, a ligação do aparelho transcricional não é possível devido à posição dos nucleossomos perto do promotor. Assim, a estrutura da cromatina geralmente deve ser mudada para ativar a transcrição eucariótica. A estrutura da cromatina ao redor de genes ativados ou reprimidos dentro das células pode ser bem estável e herdada pelas células-filhas. A herança da estrutura da cromatina é uma forma de herança que não envolve diretamente a seqüência de DNA.

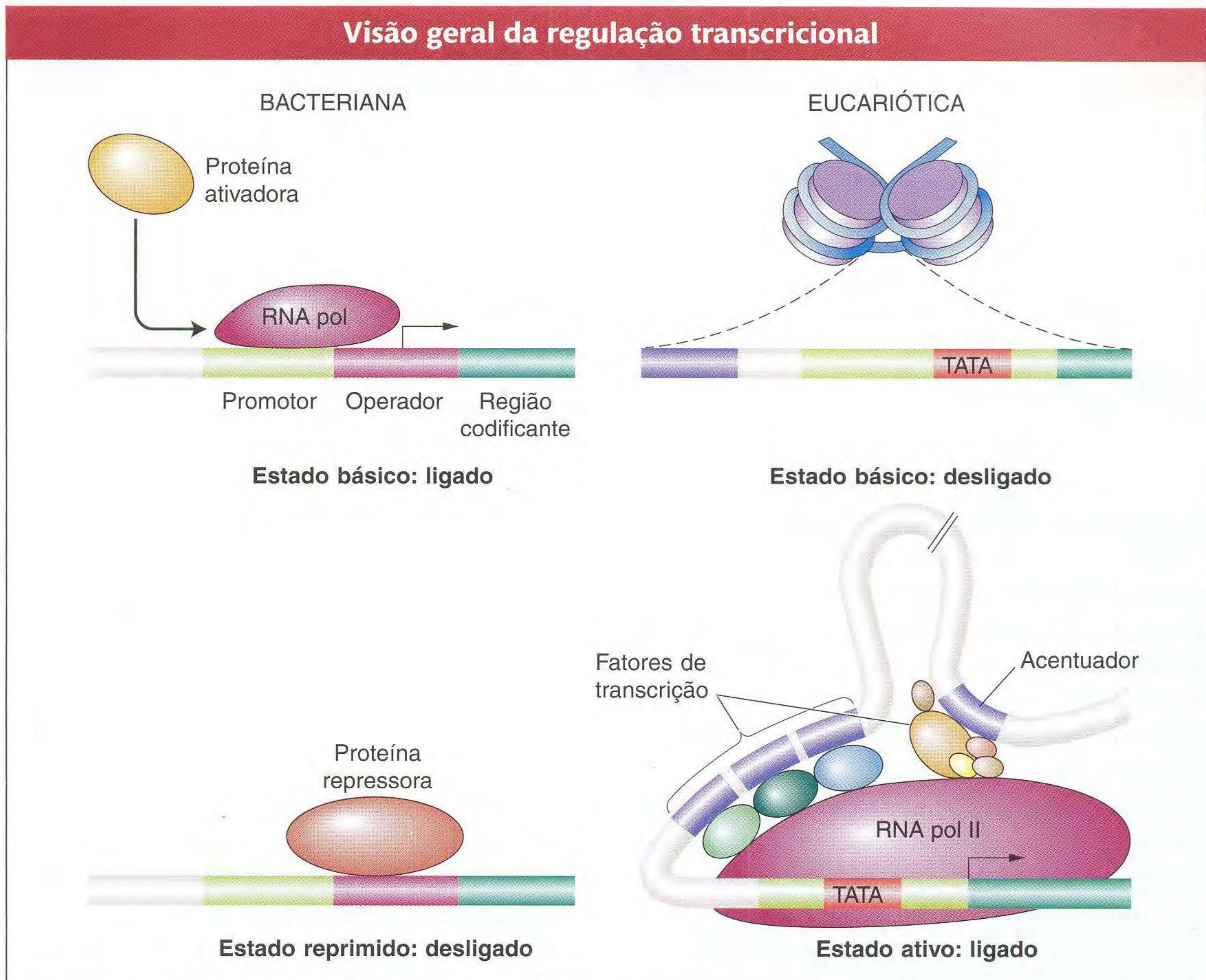


FIG. 11.2 Em bactérias, a RNA polimerase geralmente pode começar a transcrição, a menos que uma proteína repressora a bloqueie. Nos eucariontes, entretanto, a compactação do DNA com nucleossomos impede a transcrição, a menos que outras proteínas reguladoras estejam presentes. Essas proteínas reguladoras expõem as seqüências promotoras alterando a densidade dos nucleossomos ou a posição. Elas também podem recrutar a RNA polimerase II mais diretamente por ligação.

As características únicas da regulação transcricional eucariótica são o foco do restante deste capítulo. Algumas diferenças da regulação transcricional em bactérias já foram notadas no Capítulo 8:

1. Nas bactérias, todos os genes são transcritos em RNA pela mesma RNA polimerase, enquanto três RNA polimerases funcionam nos eucariontes. A RNA polimerase II, que transcreve mRNA, foi o foco do Capítulo 8 e será a única polimerase discutida neste capítulo.
2. Os RNA transcritos são amplamente processados durante a transcrição em eucariontes; as pontas 5' e 3' são modificadas, e os íntrons são removidos.
3. A RNA polimerase II é muito maior e mais complexa que sua contraparte bacteriana. Um motivo para a complexidade adicional é que a RNA polimerase II deve sintetizar RNA e coordenar os eventos únicos de processamento especial em eucariontes.

Os eucariontes multicelulares podem ter até 25.000 genes, várias vezes mais que a bactéria média. Além disso, os padrões de expressão gênica eucariótica podem ser extraordinariamente complexos. Isto é, há uma grande variação entre os genes quando um gene está ligado (transcrito) ou desligado

(não-transcrito) e em quanto do transcrito precisa ser feito. Por exemplo, um gene pode ser transcrito apenas durante o início do desenvolvimento e outro apenas na presença de uma infecção viral. Finalmente, a maioria dos genes em uma célula eucariótica estão desligados em qualquer momento. Com base apenas nessas considerações, a regulação gênica eucariótica deve ser capaz de:

1. garantir que a expressão da maioria dos genes no genoma está desligada em qualquer época, enquanto ativada em um subgrupo de genes; e
2. gerar milhares de padrões da expressão gênica.

Como você verá mais adiante neste capítulo, evoluíram mecanismos para garantir que a maioria dos genes em uma célula não são transcritos. Antes de considerar como os genes são transcricionalmente inativos, enfocaremos um segundo ponto: Como os genes eucarióticos são capazes de exibir um número enorme de diversidade de padrões de expressão? A maquinaria necessária para gerar tantos padrões de transcrição gênica *in vivo* tem muitos componentes, incluindo tanto proteínas reguladoras quanto seqüências de ação cis. O primeiro grupo de proteínas compreende um grande complexo de RNA polimerase II e os fatores gerais de transcrição que

Os elementos proximais do promotor precedem o promotor de um gene eucariótico

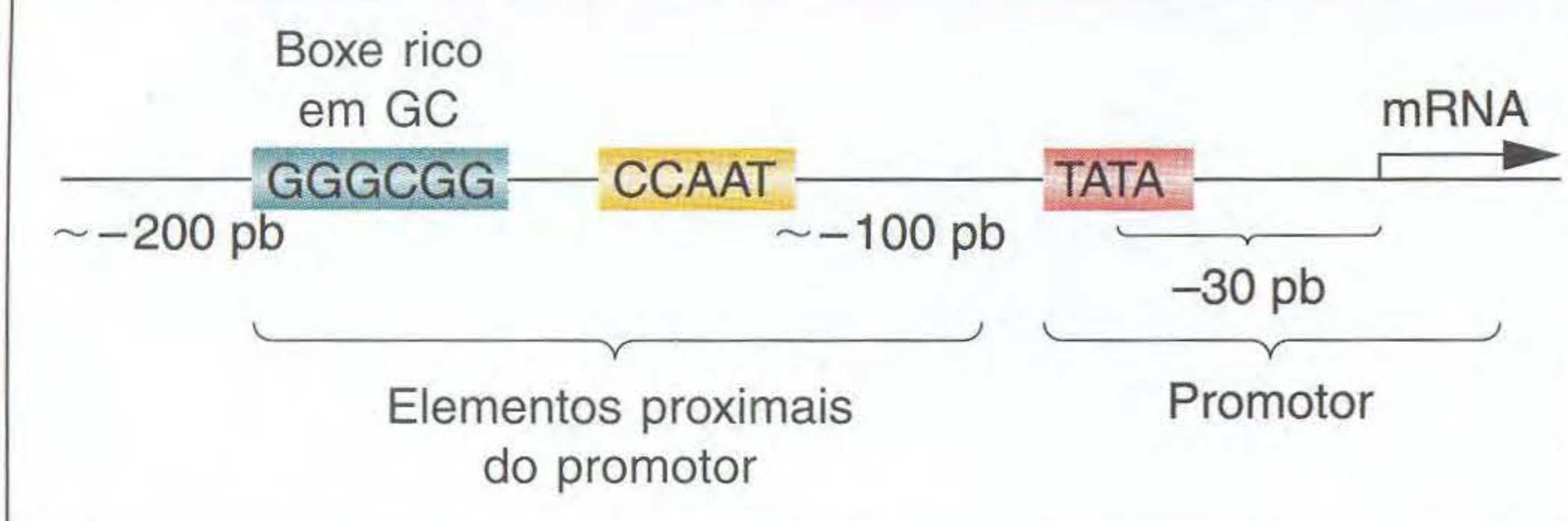


FIG. 11.3 A região antecedente do ponto de início da transcrição em eucariontes superiores contém elementos proximais do promotor e o promotor.

você aprendeu no Capítulo 8. Para iniciar a transcrição, essas proteínas interagem com seqüências de DNA chamadas **elementos proximais do promotor**, perto do promotor de um gene. O segundo grupo de componentes de proteína consiste em fatores específicos de transcrição que se ligam a seqüências reguladoras de ação cis no DNA chamados **acentuadores** ou **seqüências de ativação antecedentes (UAS)**. Essas seqüências reguladoras podem estar localizadas a uma considerável distância dos promotores gênicos. De um modo geral, os promotores e os elementos proximais dos promotores são ligados por fatores de transcrição que afetam a expressão de muitos genes. Os acentuadores são os alvos de fatores de transcrição mais específicos que controlam a regulação de subgrupos menores de genes. Geralmente, um acentuador irá atuar apenas em um ou alguns tipos de células em um eucarionte multicelular.

Para a RNA polimerase II transcrever o DNA em RNA a uma taxa máxima, vários elementos reguladores de ação cis

devem participar. Os promotores, elementos proximais dos promotores, e acentuadores são todos alvos para a ligação de diferentes proteínas de ação trans de ligação ao DNA. A Figura 11.3 é uma representação esquemática do promotor e dos elementos de seqüência proximal do promotor. A ligação da RNA polimerase II ao promotor não produz uma transcrição eficiente em si. A transcrição requer a ligação de fatores gerais de transcrição a elementos adicionais proximais do promotor que são comumente encontrados dentro de 100 pb do sítio de início da transcrição de muitos (mas não de todos) os genes. Um desses elementos é o CCAAT boxe, e outro geralmente é um segmento rico em GC mais distante antecedentemente. Os fatores gerais de transcrição que se ligam aos elementos proximais do promotor são expressos na maioria das células, e assim eles estão disponíveis para iniciar a transcrição a qualquer tempo. As mutações nesses sítios podem ter um marcante efeito na transcrição, demonstrando o quanto eles são importantes. Um exemplo das conseqüências nas taxas de transcrição de mutações nesses elementos de seqüência é mostrado na Figura 11.4.

Para modularem a transcrição, as proteínas reguladoras possuem um ou mais dos seguintes *domínios funcionais*:

1. Um domínio que reconhece uma seqüência reguladora no DNA (o sítio de ligação da proteína ao DNA)
2. Um domínio que interage com uma ou mais proteínas do aparelho de transcrição (RNA polimerase ou uma proteína associada à RNA polimerase)
3. Um domínio que interage com as proteínas ligadas a seqüências reguladoras vizinhas no DNA de tal modo que atuam cooperativamente para regular a transcrição
4. Um domínio que influencia a condensação da cromatina seja direta ou indiretamente

Os elementos proximais do promotor são necessários para transcrição eficiente

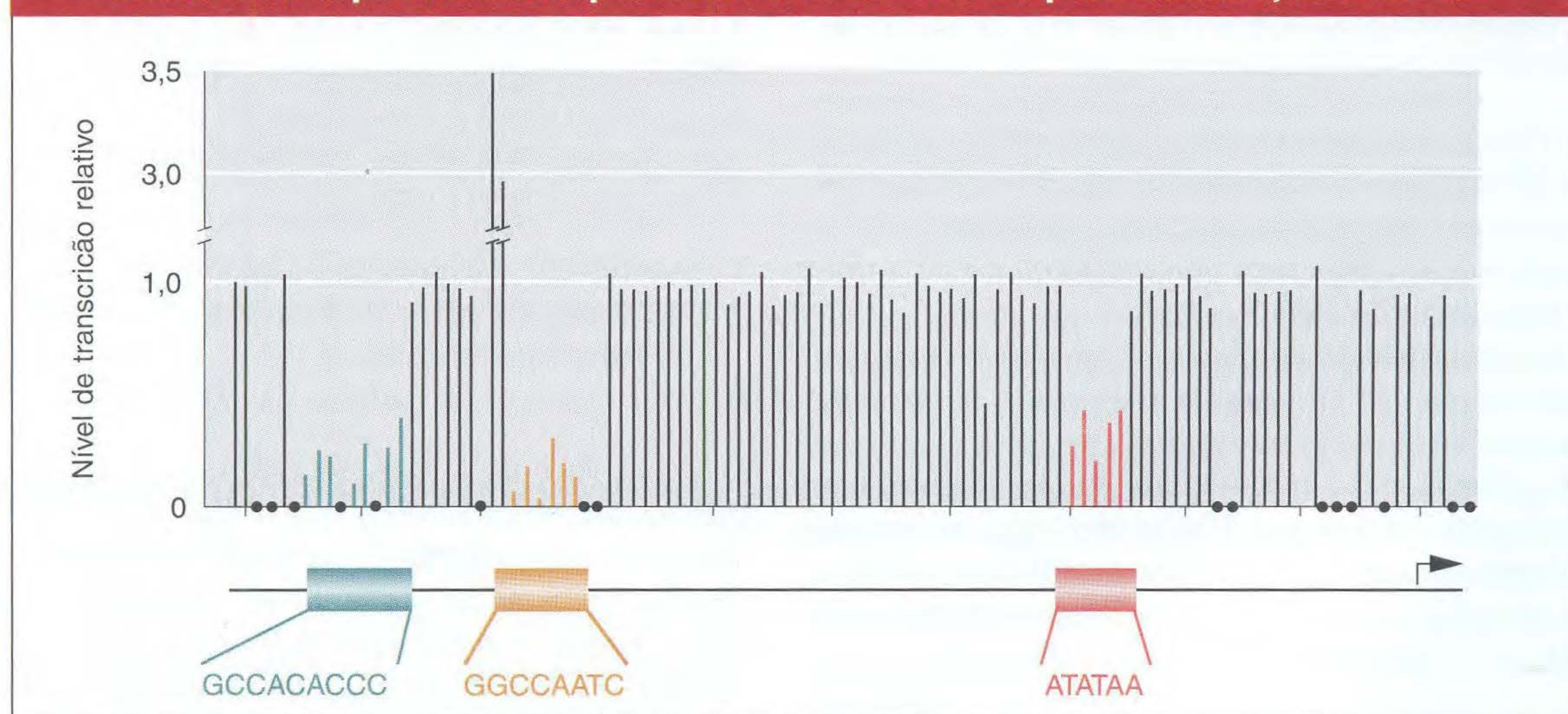


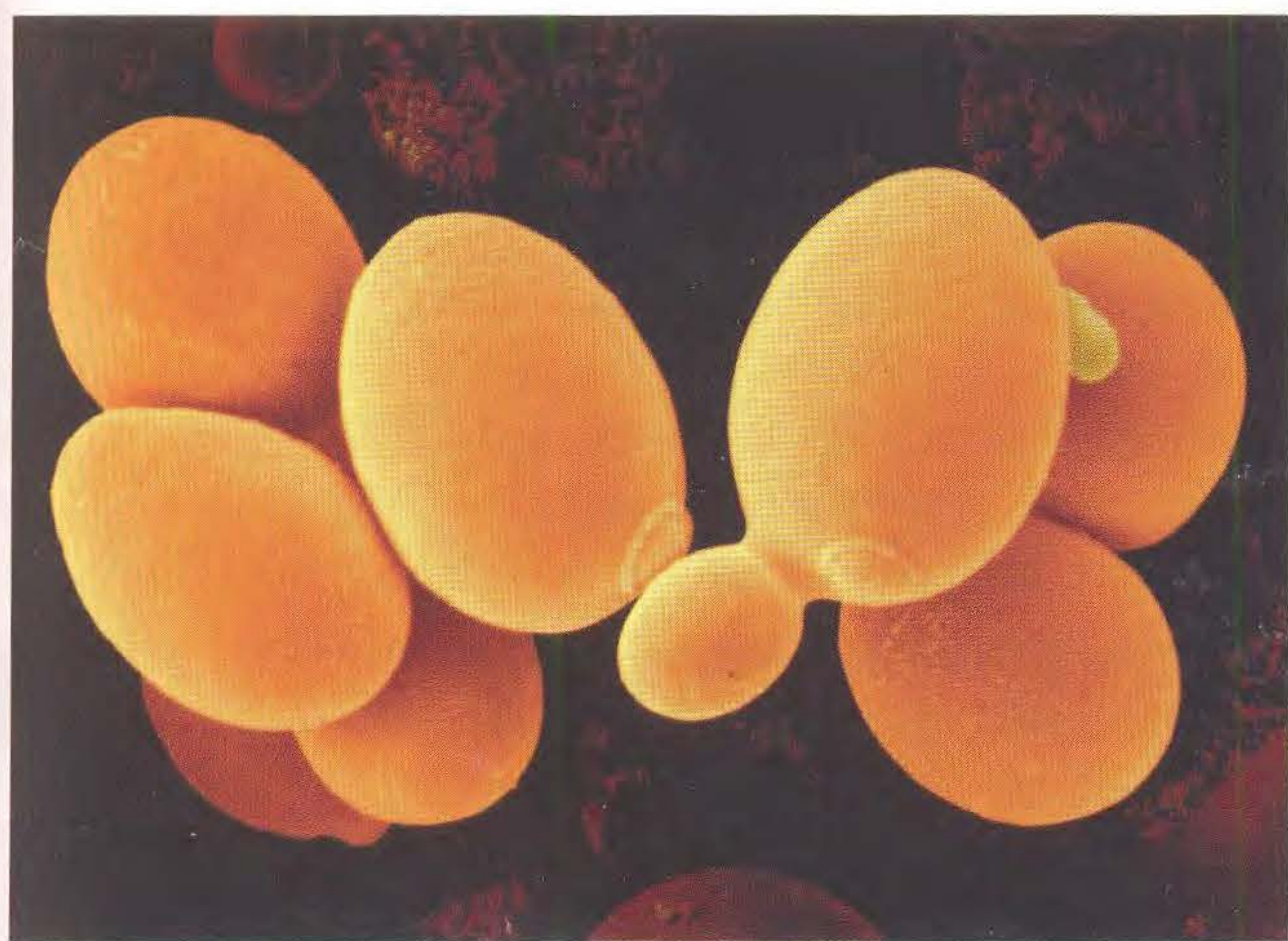
FIG. 11.4 As mutações de ponto no promotor e elementos proximais do promotor impedem a transcrição do gene de β -globina. As mutações de ponto através da região promotora foram analisadas por seus efeitos nas taxas de transcrição. A altura de cada linha representa o nível de transcrição relativo a um promotor tipo selvagem ou elemento proximal do promotor (1,0). Apenas uma substituição de bases que fica dentro dos três elementos mostrados muda o nível de transcrição. As posições com pontos pretos não foram testadas. [De T. Maniatis, S. Goodbourn, and J. A. Fischer, "Regulation of Inducible and Tissue-Specific Gene Expression," *Science* 236, 1987, 1237.]



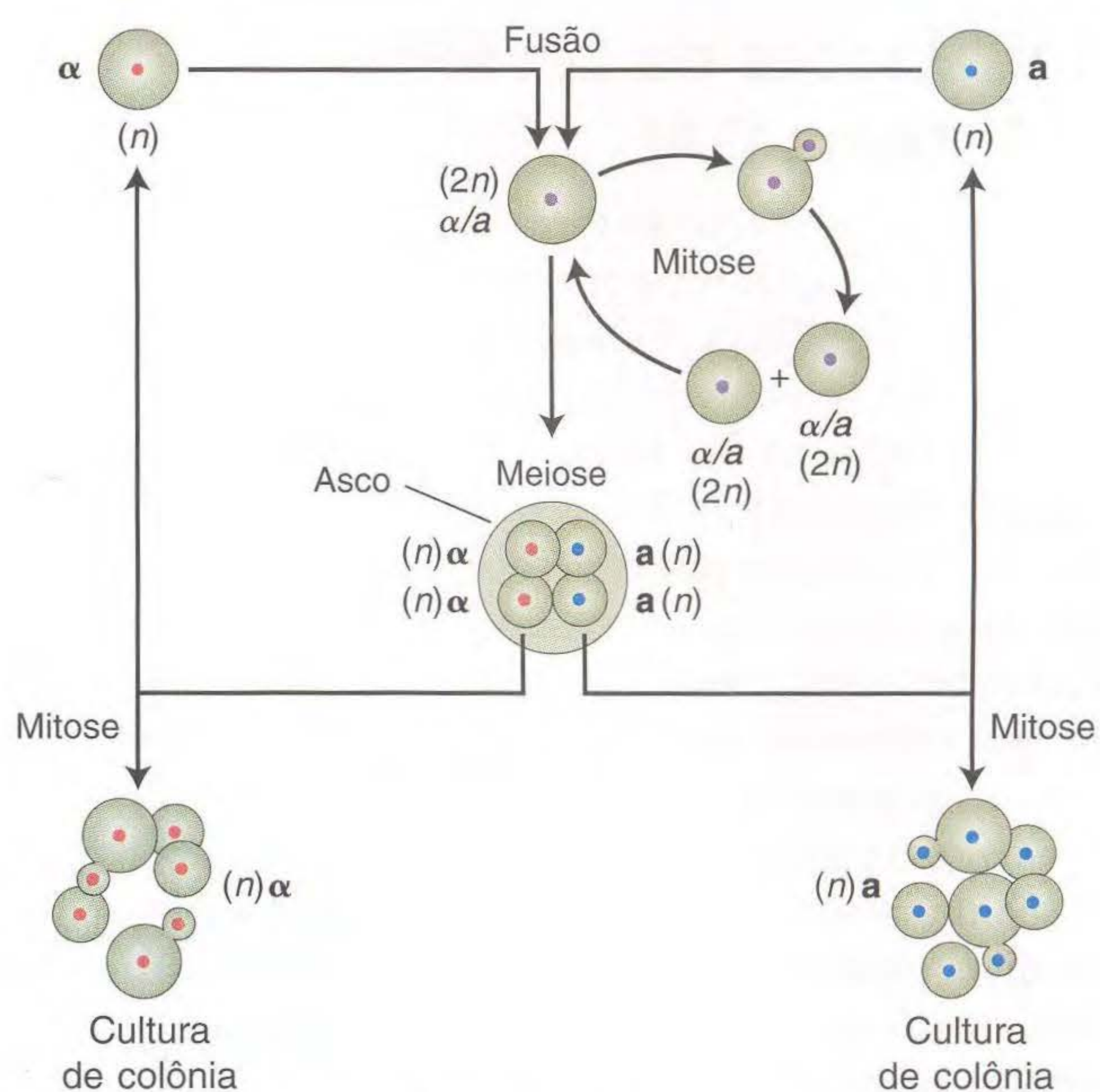
Organismo-modelo Leveduras

Saccharomyces cerevisiae, ou levedura de brotamento, surgiu em anos recentes como um importante sistema genético eucariótico. Os humanos têm cultivado leveduras por séculos porque é um componente essencial da cerveja, pão e vinho. As leveduras têm muitas características que as tornam um organismo-modelo ideal. Como um eucarionte unicelular, ela pode crescer em placas de ágar e, com um ciclo vital apenas de 90 min, grandes quantidades podem ser cultivadas em meio líquido. Ela tem um genoma muito compacto apenas com 12 pares de megabases de DNA (comparada com quase 3.000 pares de megabases em humanos) contendo aproximadamente 6.000 genes que são distribuídos em 16 cromossomos. Foi o primeiro eucariote a ter seu genoma seqüenciado.

O ciclo de vida das leveduras as torna muito versáteis para estudos de laboratório. As células podem crescer como diplóides ou haplóides. Em ambos os casos, a célula-mãe produz um brotamento contendo uma célula-filha idêntica. As células diplóides ou continuam crescendo por brotamento, ou são induzidas a sofrer meiose, que produz quatro esporos haplóides mantidos juntos em um *asco* (também chamado de *tétrade*). Os esporos haplóides de tipo reprodutivo oposto (**a** ou α) irão fundir-se e formar um diplóide. Os esporos do mesmo tipo reprodutivo continuarão a crescer por brotamento.



Micrografia eletrônica de células de levedura em brotamento.
[SciMAT/Photo Researchers.]



O ciclo de vida da levedura do pão. Os alelos nucleares *MATa* e *MATα* determinam o tipo reprodutivo.

As leveduras têm sido chamadas de *E. coli* dos eucariontes devido à facilidade de análise de mutantes diretos e reversos. Para isolar os mutantes usando um enfoque genético direto, as células haplóides são mutagenizadas (com raios X, por exemplo) e triadas em placas para fenótipos mutantes. Esse procedimento geralmente é feito primeiro plaqueando células em meio rico no qual todas as células crescem e copiando, ou *plaqueamento réplica*, as colônias a partir desta placa master em placas réplica que contêm o meio seletivo ou as condições especiais de crescimento. (Veja também o Capítulo 15.) Por exemplo, os mutantes sensíveis à temperatura crescerão na placa master na temperatura permissiva, mas não na placa réplica em temperatura restritiva. As comparações das colônias nas placas master e réplica revelarão os mutantes sensíveis à temperatura. Usando genética reversa, os cientistas também podem substituir qualquer gene de levedura (de função conhecida ou desconhecida) por uma versão mutante (sintetizada em um tubo de ensaio) para compreender a natureza do produto gênico.

- Um domínio que atua como um sensor de condições fisiológicas dentro da célula.

Grande parte da estratégia do controle transcricional eucariótico é baseada em como os fatores específicos de transcrição controlam o acesso dos fatores gerais de transcrição e a RNA polimerase II. Os mecanismos reguladores de genes eucarióticos foram descobertos por enfoques tanto

bioquímicos quanto genéticos. Estes últimos foram desenvolvidos em particular por estudos na levedura unicelular *Saccharomyces cerevisiae* (veja o boxe Organismo-modelo *Leveduras*). Várias décadas de pesquisa foram uma fonte de muitos *insights* sobre os princípios gerais de como as proteínas reguladoras da transcrição eucariótica funcionam, e de como diferentes tipos celulares são gerados. Examinaremos em detalhe dois sistemas reguladores de genes: o primeiro é

relacionado à via de uso de galactose; o segundo é o controle do tipo reprodutivo.

► 11.2 Lições de Leveduras: o Sistema GAL

Para fazer uso de galactose extracelular, as leveduras importam o açúcar e o convertem em uma forma de glicose que pode ser metabolizada. Vários genes — *GAL1*, *GAL2*, *GAL7* e *GAL10* — no genoma de leveduras codificam enzimas que catalisam etapas na via bioquímica, convertendo galactose em glicose (Figura 11.5). Três genes adicionais — *GAL3*, *GAL4* e *GAL80* — codificam proteínas que regulam a expressão dos genes de enzimas. Como no sistema *lac*, a abundância de açúcar determina o nível de expressão gênica na via bioquímica. Em células de leveduras que crescem em meios sem galactose, os genes *GAL* estão amplamente silenciados. Mas, na presença de galactose (e na ausência de glicose), os genes *GAL* são induzidos. Do mesmo modo que para o óperon *lac*, as análises genéticas e moleculares dos mutantes têm sido fundamentais para compreender como a expressão dos genes na via de galactose é controlada.

O principal regulador da expressão gênica de *GAL* é a proteína Gal4, uma proteína de ligação ao DNA específica de sequência. Gal4 é talvez a proteína reguladora transcricional mais bem estudada em eucariontes. A dissecção detalhada de sua regulação e atividade foi uma fonte de vários *insights* importantes no controle da transcrição em eucariontes.

Gal4 regula vários genes pelas seqüências de ativação antecedentes

Na presença de galactose, os genes *GAL1*, *GAL2*, *GAL7* e *GAL10* são induzidos 1.000 vezes ou mais. Nos mutantes *GAL4*, entretanto, eles permanecem silenciosos. Cada um desses quatro genes tem dois ou mais sítios de ligação Gal4 situados em 5' (antecedentes) de seu promotor. Considere os genes *GAL10* e *GAL1*, que são adjacentes um ao outro e trans-

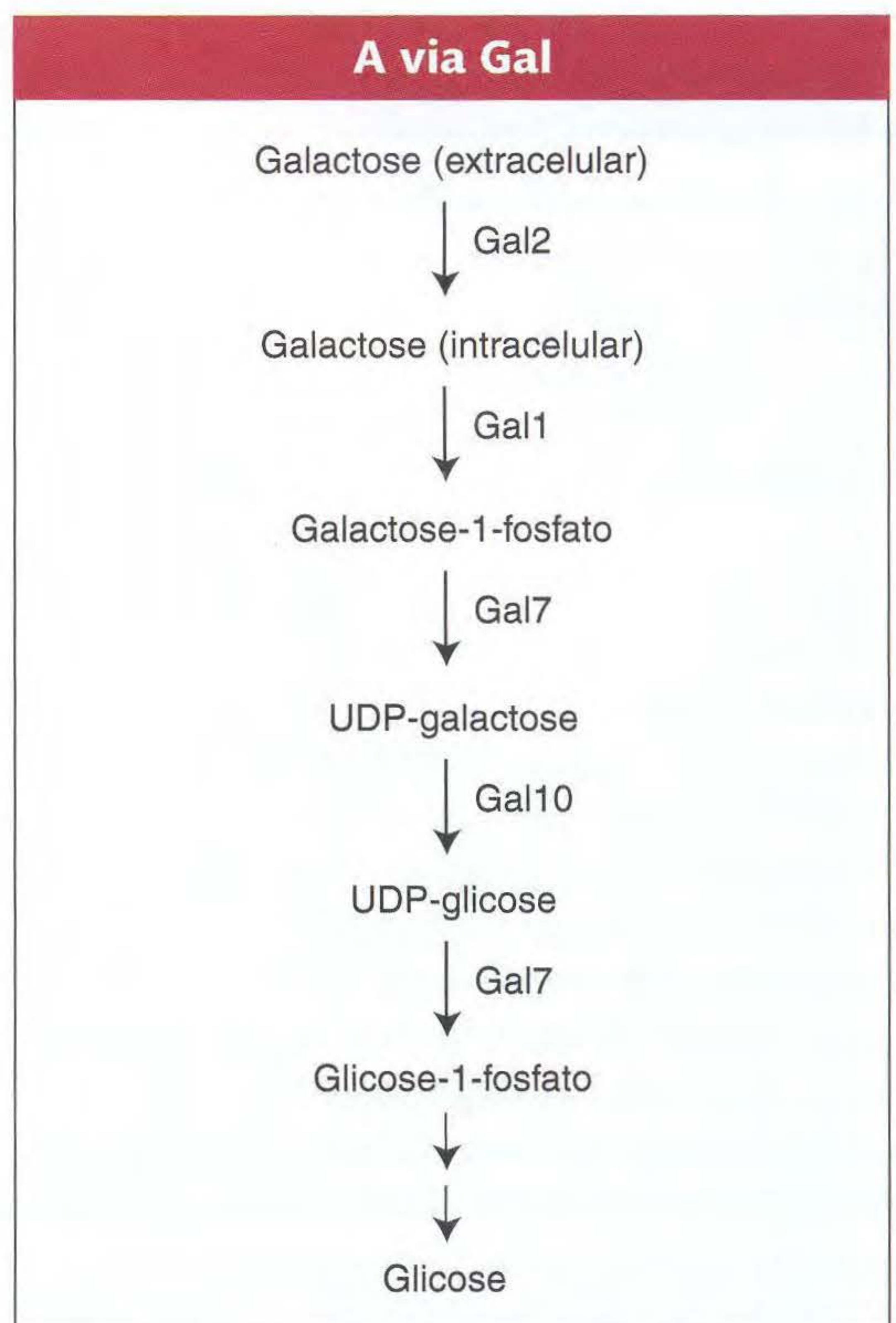


FIG. 11.5 A galactose é convertida em glicose-1-fosfato em uma série de etapas. Essas etapas são catalisadas por enzimas (Gal1 e assim por diante) codificadas por genes estruturais *GAL1*, *GAL2*, *GAL7* e *GAL10*.

critos em sentidos opostos. Entre o sítio de início da transcrição de *GAL1* e o sítio do início de transcrição de *GAL10*, há uma única região de 118 pb que contém quatro sítios de ligação Gal4 (Figura 11.6). Cada sítio de ligação Gal4 tem 17 pares de bases de tamanho e é ligado por um dímero de proteína Gal4. Existem dois sítios de ligação Gal4 antecedentes ao gene *GAL2*, e outros dois antecedentes ao gene *GAL7*. Esses sítios de ligação são necessários para a ativação gênica *in vivo*. Se eles forem deletados, os genes são silenciosos, mes-

As proteínas ativadoras transcricionais ligam-se a elementos UAS em levedura

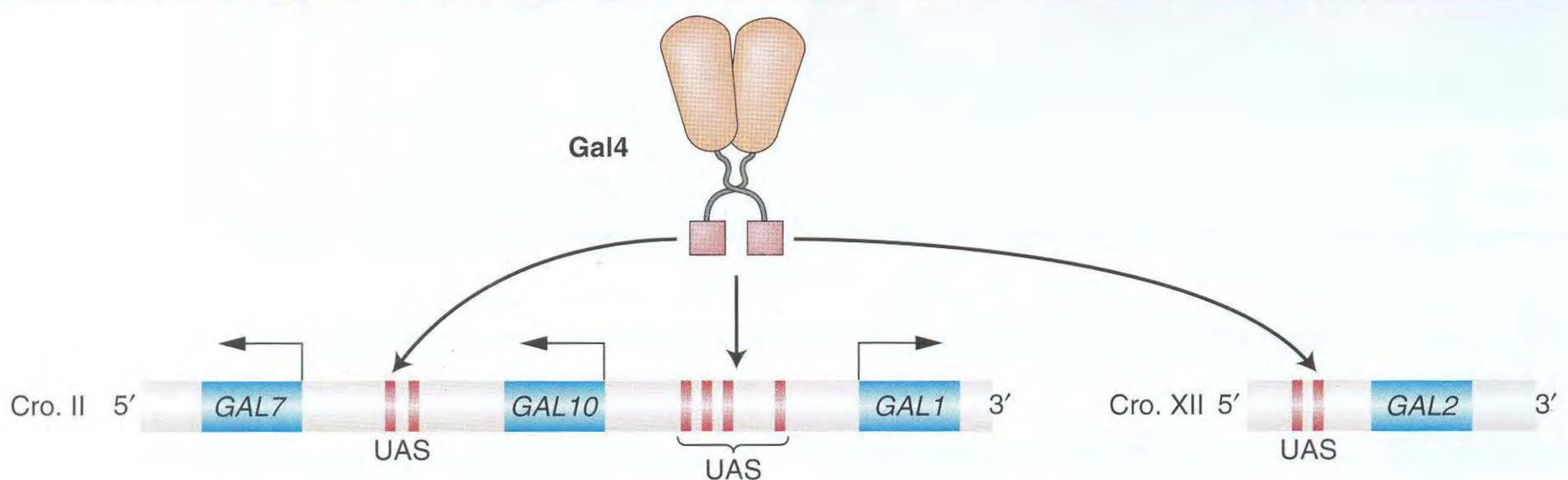


FIG. 11.6 A proteína Gal4 ativa genes-alvo pelos elementos da seqüência de ativação antecedente (UAS). A proteína Gal4 tem dois domínios funcionais: o domínio de ligação ao DNA (quadrados vermelhos) e um domínio de ativação (oval laranja). A proteína liga-se a seqüências específicas antecedentes dos promotores dos genes da via Gal. Alguns dos genes *GAL* são adjacentes (*GAL1*, *GAL10*), enquanto outros estão em cromossomos diferentes. O elemento UAS *GAL1* contém quatro sítios de ligação Gal4.

mo na presença de galactose. Essas seqüências reguladoras são acentuadores que também são chamados de seqüências de ativação antecedentes. A presença de acentuadores situados a uma distância linear considerável do promotor de um gene eucariótico é típica.

A união de proteínas de ligação ao DNA específicas de seqüências a regiões fora dos promotores de genes-alvo é uma característica comum da regulação transcricional eucariótica.

A proteína Gal4 tem domínios de ligação ao DNA e ativação separáveis

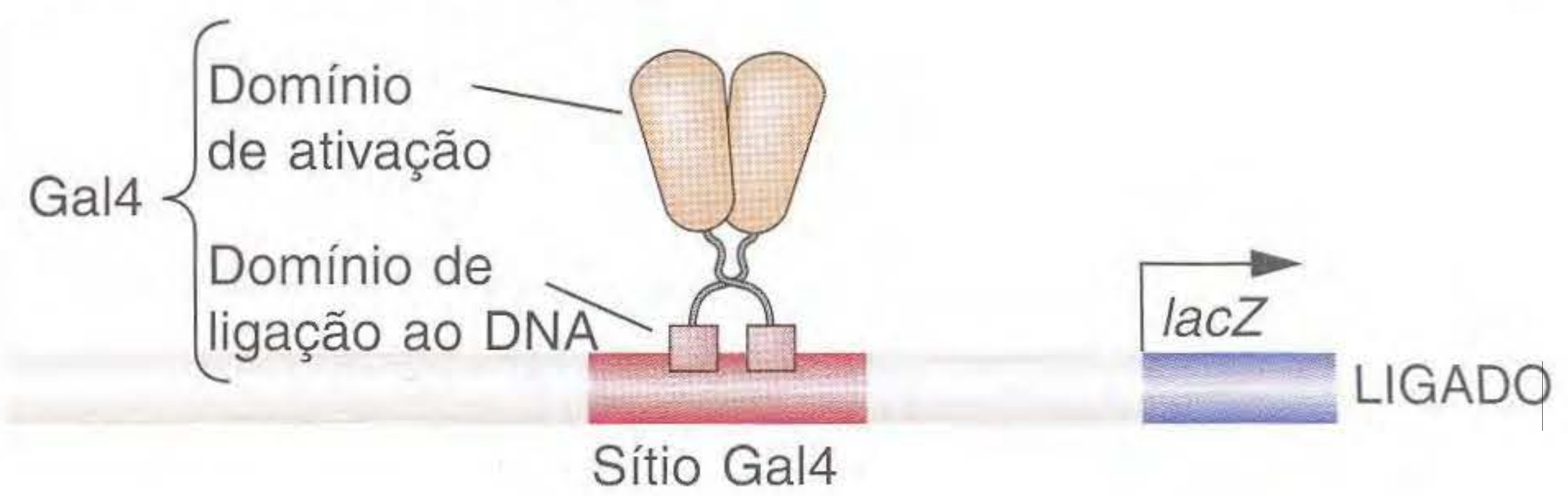
Após Gal4 estar ligado ao elemento UAS, como é induzida a expressão do gene? Um domínio distinto da proteína Gal4, o **domínio de ativação**, é necessário para a atividade reguladora. Assim, a proteína Gal4 tem pelo menos dois domínios: um para a ligação ao DNA e outro para ativação da transcrição. Uma organização modular similar foi encontrada como sendo uma característica comum também de outros fatores de transcrição de ligação ao DNA.

A organização modular da proteína Gal4 foi demonstrada em uma série de experimentos simples e elegantes. A estratégia foi testar a ligação ao DNA e a ativação gênica de formas mutantes da proteína nas quais partes haviam sido deletadas ou fundidas a outras proteínas. Desse modo, podia ser determinado se uma parte da proteína era necessária para determinada função. Para fazer esses estudos, os experimentadores precisaram de um meio simples de dosar a expressão das enzimas codificadas pelos genes *GAL*. A expressão dos genes *GAL* e de outros alvos de fatores de transcrição é tipicamente monitorada usando um **gene repórter** cuja expressão é facilmente rastreada. Em geral, o gene repórter é o gene *lacZ* de *E. coli*, que pode atuar em substratos cujos produtos são facilmente avaliados por sua cor brilhante ou fluorescência. Outro gene repórter comum é o gene que codifica a proteína fluorescente verde da água-viva, que, como o nome sugere, é facilmente rastreada pela luz que ela emite. A região codificante de um desses genes repórter e um promotor são colocados posteriores a um elemento UAS do gene *GAL*. A expressão do repórter é então lida da atividade Gal4 nas células.

Quando uma forma da proteína Gal4 sem o domínio de ativação é expressa em leveduras, os sítios de ligação do elemento UAS são ocupados, mas nenhuma transcrição é estimulada (Figura 11.7). O mesmo é verdadeiro quando outras proteínas reguladoras sem domínios de ativação, tais como o repressor bacteriano LexA, são expressas nas células que possuem genes repórter com seus respectivos sítios de ativação. O resultado mais interessante é obtido quando uma forma da proteína Gal4 sem o domínio de ligação ao DNA é inserida no domínio de ligação ao DNA LexA; a proteína híbrida agora ativa a transcrição dos sítios de ligação de LexA (Figura 11.7). Experimentos posteriores de “troca de domínio” revelaram que a ativação funcional

As proteínas ativadoras transcricionais são modulares

(a) O dímero Gal4 completo



(b) Gal4 sem o domínio de ativação



(c) LexA sem domínio de ativação



(d) Híbrido Gal4-LexA

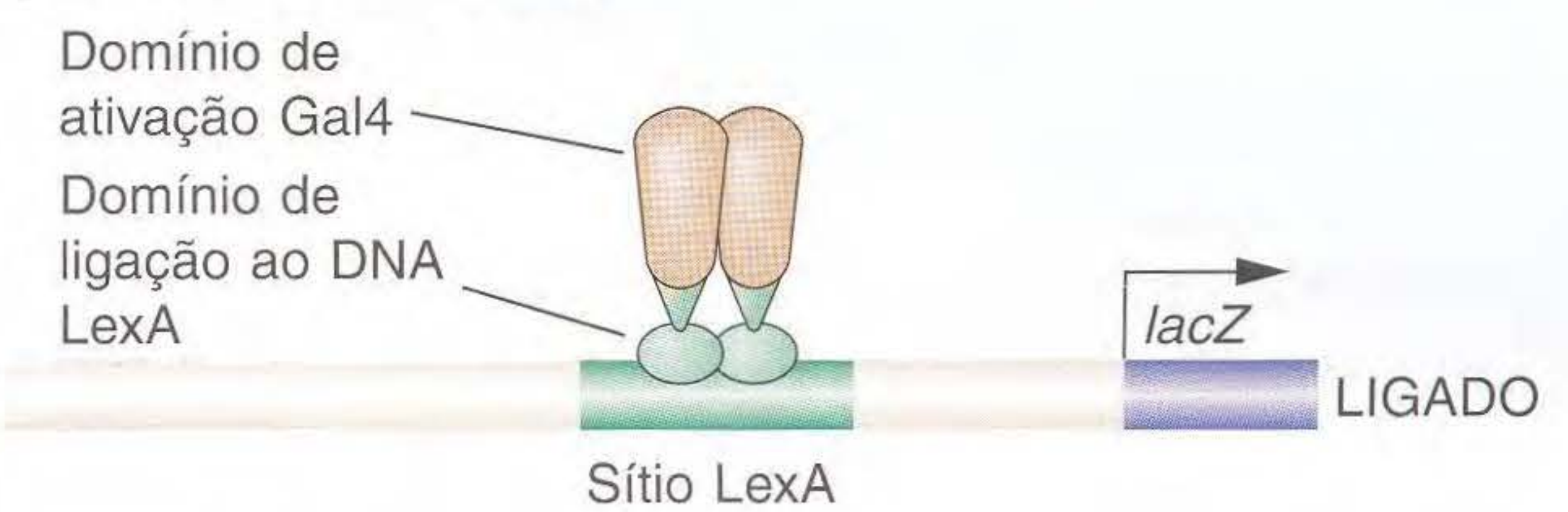


FIG. 11.7 As proteínas ativadoras transcricionais têm vários domínios separáveis. (a) A proteína Gal4 tem dois domínios e forma um dímero. (b) A remoção experimental do domínio de ativação mostra que a ligação do DNA não é suficiente para a ativação gênica. (c) Similarmente, a proteína LexA bacteriana não pode ativar a transcrição por si, mas, quando fundida ao domínio de ativação Gal4 (d), pode ativar a transcrição pela ligação com os sítios de ligação LexA. [De J. Watson et al., *Molecular Biology of the Gene*, Quinta Edição, copyright © 2004, Benjamin Cummings.]

transcricional da proteína Gal4 reside em dois pequenos domínios com cerca de 50 a 100 aminoácidos de tamanho. Esses domínios são separáveis dos usados na dimerização da proteína, ligação ao DNA e interação com a proteína Gal80 (veja em seguida). O domínio de ativação ajuda a recrutar a maquinaria transcricional para o promotor, como veremos na Seção 11.3. Esse arranjo altamente modular de domínios de regulação de atividade é encontrado em muitos fatores de transcrição.

Muitas proteínas eucarióticas reguladoras de transcrição são proteínas modulares, com domínios separáveis de ligação ao DNA, ativação ou repressão e interação com outras proteínas.

A atividade de Gal4 é regulada fisiologicamente

Como Gal4 torna-se ativa na presença de galactose? Indícios-chave vieram das análises de mutações nos genes *GAL80* e *GAL3*. Nos mutantes *GAL80*, os genes estruturais *GAL* são ativos mesmo na ausência de galactose. Esse resultado sugere que a função normal da proteína Gal80 é, de algum modo, inibir a expressão do gene *GAL*. Contrariamente, nos mutantes *GAL3*, os genes estruturais *GAL* não são ativos na presença de galactose, sugerindo que Gal3 promove normalmente a expressão dos genes *GAL*.

Amplas análises bioquímicas revelaram que a proteína Gal80 liga-se à proteína Gal4 com alta afinidade e inibe diretamente a atividade de Gal4. Especificamente, Gal80 liga-se a uma região dentro de um dos domínios de ativação de Gal4, bloqueando sua habilidade em promover a transcrição dos genes-alvo. O papel das proteínas Gal3 é liberar Gal4 de sua inibição por Gal80 na presença de galactose. Gal3 é um sensor e indutor. Quando Gal3 se liga a galactose e ATP, ela sofre uma mudança alostérica que promove a ligação a Gal80, que, por sua vez, faz com que Gal80 libere Gal4, e esta é, então, capaz de ativar a transcrição de seus genes-alvo. Assim, Gal3, Gal80 e Gal4 são todas partes de uma mudança cujo estado é determinado pela presença ou ausência de galactose (Figura 11.8). Nessa mudança, a ligação ao DNA pelo regulador transcricional não é a etapa regulada fisiologicamente (como é o caso no óperon *lac* e bacteriófago λ). A atividade do domínio de ativação é regulada.

A atividade das proteínas reguladoras transcricionais eucarióticas é geralmente controlada pelas interações com outras proteínas.

Funções de Gal4 na maioria dos eucariontes

Em adição à sua ação nas células de leveduras, Gal4 tem sido demonstrada como sendo capaz de ativar a transcrição em células de insetos, células humanas e muitas outras espécies eucarióticas. Essa versatilidade sugere que a maquinaria bioquímica e os mecanismos de ativação gênica são comuns a uma ampla gama de eucariontes, e que as características reveladas em leveduras estão geralmente presentes em outros eucariontes, e vice-versa. Além disso, devido à sua versatilidade, Gal4 e seus elementos UAS tornaram-se as ferramentas favoritas na análise genética para manipular a expressão gênica e função em uma grande variedade de sistemas-modelo.

A habilidade de Gal4, assim como de outros reguladores eucarióticos, em funcionar em uma variedade de eucariontes em geral tem a maquinaria reguladora transcricional e os mecanismos em comum.

Agora veremos como os ativadores e outras proteínas reguladoras interagem com a maquinaria transcricional para controlar a expressão gênica.

Ativadores recrutam a maquinaria transcricional

Nas bactérias, os ativadores comumente estimulam a transcrição interagindo diretamente com o DNA e com a RNA polimerase. Em eucariontes, os ativadores geralmente trabalham indiretamente para recrutar a RNA polimerase II para genes promotores por meio de dois mecanismos importantes. Primeiro, os ativadores podem interagir com subunidades dos complexos proteicos, tendo papéis no início da transcrição. Segundo, os ativadores podem recrutar proteínas que modificam a estrutura da cromatina, permitindo que a RNA polimerase II e outras proteínas tenham acesso ao DNA. Muitos ativadores, incluindo Gal4, têm ambas as atividades. Examinaremos primeiro o recrutamento de partes do complexo de iniciação transcricional.

Lembre, do Capítulo 8, que a maquinaria transcricional eucariótica contém muitas proteínas que são partes de vários subcomplexos dentro do aparato transcricional montado nos promotores gênicos. Um subcomplexo, o fator de transcrição IID (TFIID), liga-se ao TATA box dos promotores eucarióticos pela proteína de ligação a TATA (TBP; veja Figura 8.12). Gal4 liga-se à TBP em um sítio de seu domínio de ativação, e, por essa ligação, recruta o complexo TFIID e, por sua vez, a RNA polimerase II ao promotor (Figura 11.9). A afinidade dessa interação correlaciona-se bem com a potência de Gal4 como um ativador. Gal4 também interage com o grande **complexo mediador**, que interage diretamente com a RNA polimerase II para recrutá-la para os promotores gênicos. O complexo mediador é um exemplo de um **coativador**, um

As proteínas ativadoras transcricionais podem ser ativadas por um indutor

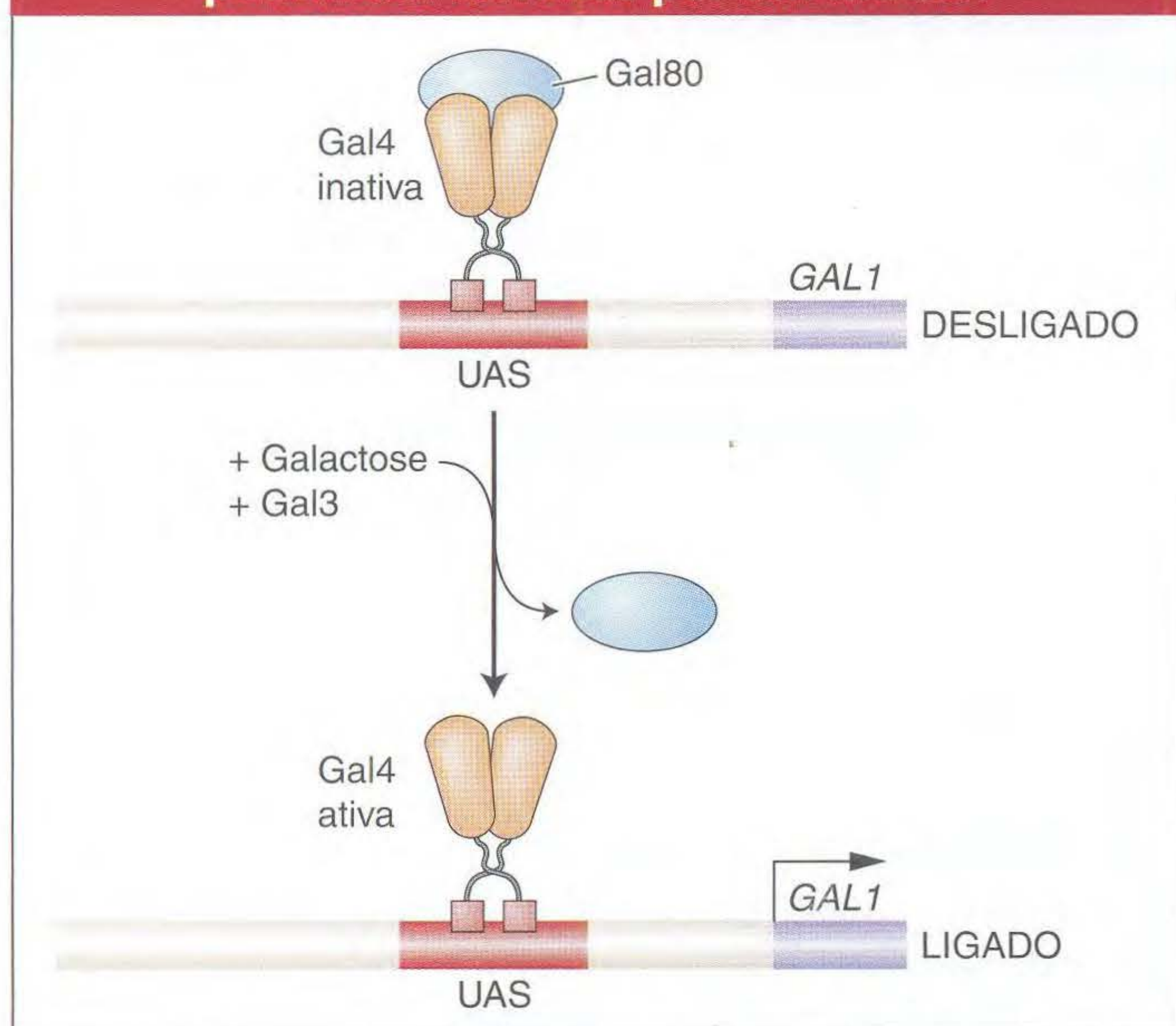


FIG. 11.8 A atividade de Gal4 é regulada pela proteína Gal80. (Em cima) Na ausência de galactose, a proteína Gal4 é inativa, muito embora possa se ligar a sítios antecedentes ao gene-alvo *GAL1*. A atividade de Gal4 é suprimida pela ligação da proteína Gal80. (Embaixo) Na presença de galactose e proteína Gal3, Gal80 sofre uma mudança conformacional e libera o domínio de ativação de Gal4, permitindo a transcrição do gene-alvo.

As proteínas ativadoras transcricionais recrutam a maquinaria transcricional

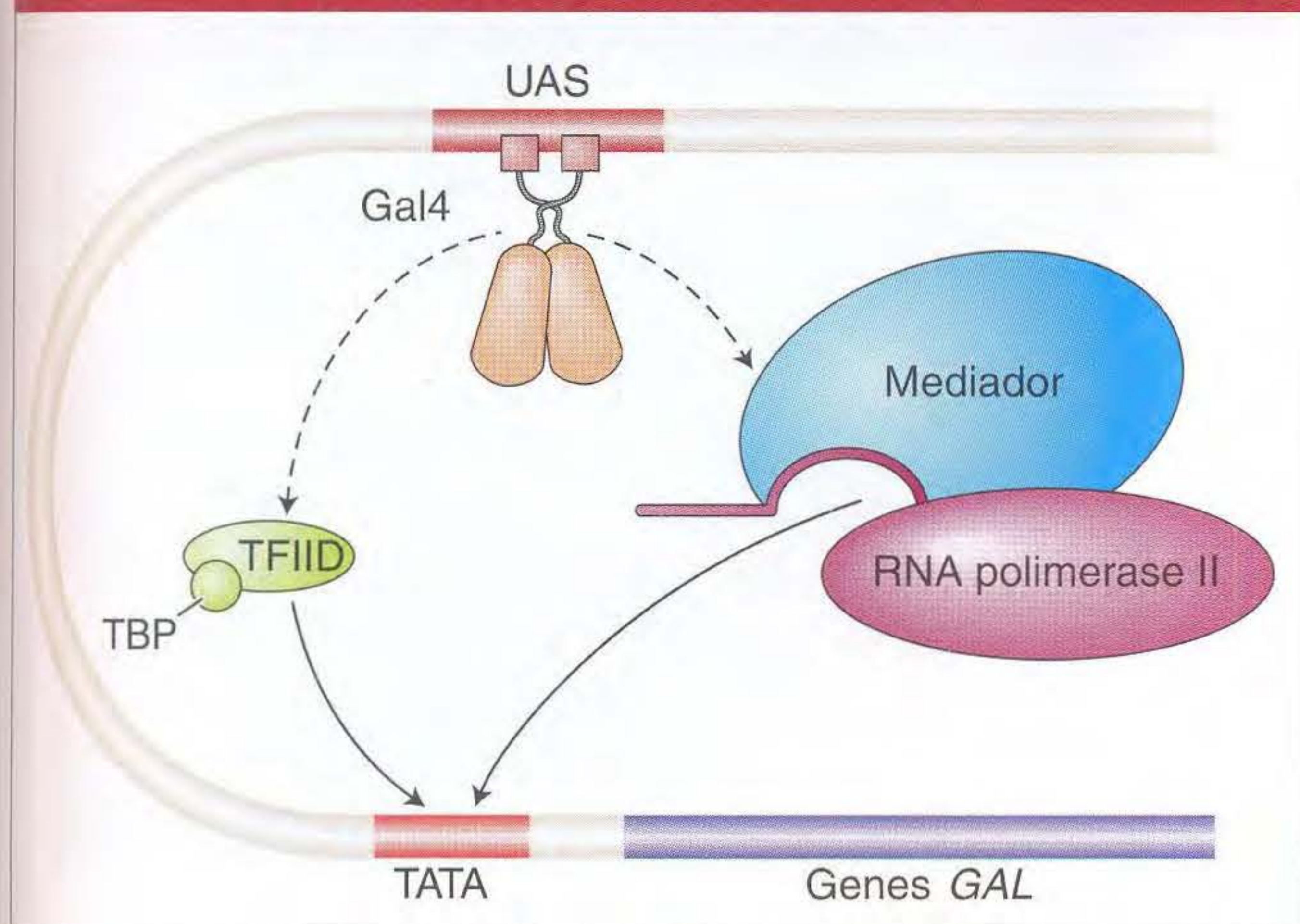


FIG. 11.9 Gal4 recruta a maquinaria transcricional. A proteína Gal4, bem como muitos outros ativadores transcricionais, liga-se a múltiplos complexos proteicos, incluindo TFIID e complexos mediadores, que recrutam a RNA polimerase II para promotores do gene. As interações facilitam a ativação gênica pelos sítios de ligação que estão distantes dos genes promotores. [De J. Watson et al., *Molecular Biology of the Gene*, Quinta Edição, copyright © 2004, Benjamin Cummings.]

termo aplicado a uma proteína ou complexo proteico que facilita a ativação gênica por um fator de transcrição mas que, em si, não é parte da maquinaria de transcrição nem de uma proteína de ligação ao DNA.

A habilidade dos ativadores em se ligar a seqüências antecedentes de DNA e interagir com proteínas que se ligam direta ou indiretamente a promotores ajuda a explicar como a transcrição pode ser estimulada por seqüências reguladoras mais distantes (veja Figura 11.9).

Os ativadores transcricionais eucarióticos geralmente funcionam recrutando partes da maquinaria transcricional para promotores gênicos.

11.3 Cromatina Dinâmica e Regulação Gênica Eucariótica

Um segundo mecanismo para influenciar a transcrição gênica em eucariontes modifica a estrutura local da cromatina ao redor de seqüências reguladoras gênicas. Para compreender totalmente como funciona esse mecanismo, precisamos primeiro rever a estrutura da cromatina e, então, considerar como ela pode mudar e como suas mudanças afetam a expressão gênica.

O recrutamento da maquinaria transcricional por ativadores pode parecer um tanto similar em eucariontes e bactérias, sendo a principal diferença o número de proteínas que interagem na maquinaria transcricional. De fato, há menos de uma década, muitos biólogos interpretavam a regulação eucariótica simplesmente como uma versão

bioquimicamente mais complicada do que havia sido descoberto em bactérias. Entretanto, essa visão mudou acentuadamente à medida que os biólogos consideraram o efeito da organização do DNA genômico nos eucariontes.

Comparado ao DNA eucariótico, o DNA bacteriano é relativamente “nu”, tornando-se prontamente acessível à RNA polimerase. Em contraste, os cromossomos eucarióticos são embalados em cromatina, que é composta de DNA e proteínas (principalmente histonas). Como mencionado resumidamente no Capítulo 2, a unidade básica da cromatina é o nucleossomo, contendo cerca de 150 pb de DNA envolvido em duas voltas ao redor de um octâmero de histonas (Figura 11.10). O octâmero de histonas é composto de duas subunidades de cada das quatro histonas: histona 2A, 2B, 3 e 4. Os nucleossomos podem se associar em estruturas de ordens superiores que condensam mais o DNA. A embalagem do DNA eucariótico em cromatina significa que a grande parte do DNA não está prontamente acessível a proteínas reguladoras e ao aparato de transcrição. Assim, enquanto os genes procarióticos estão geralmente acessíveis e “ligados”, a menos que reprimidos, os genes eucarióticos estão inacessíveis e “desligados”, a menos que ativados. Portanto, a modificação da estrutura da cromatina é uma característica distinta da regulação gênica eucariótica.

Podemos imaginar vários modos para alterar a estrutura da cromatina. Por exemplo, um mecanismo pode ser simplesmente movendo o octâmero de histona ao longo do DNA. Nos anos de 1980, foram desenvolvidas técnicas bioquímicas que permitiram aos pesquisadores determinar a posição dos nucleossomos no e ao redor de genes específicos. Nesses estudos, a cromatina foi isolada dos tecidos ou células nas quais um gene estava ligado e comparada com a cromatina do tecido onde o mesmo gene estava desligado. O resultado da maioria dos genes analisados foi que as posições dos nucleossomos mudaram, especialmente nas regiões reguladoras dos genes. Assim, as regiões do DNA que estavam enroladas em nucleossomos podiam mudar, e as posições dos nucleossomos podiam mudar no DNA de célula a célula e durante o ciclo de vida de um organismo. A transcrição pode ser reprimida quando o promotor e as seqüências flangeadoras estão enrolados em um nucleossomo e inacessíveis à RNA polimerase II. A ativação da transcrição iria, portanto, requerer o bloqueio do nucleossomo para ser reorganizado afastando as histonas ou removendo-as inteiramente. Contrariamente, quando a repressão gênica é necessária, os octâmeros de histona podem mudar para uma posição que evita a transcrição. A mudança da posição dos nucleossomos é chamada de **remodelagem de cromatina**. Agora, a remodelagem de cromatina é conhecida como sendo uma parte integrante da expressão gênica eucariótica, e grandes avanços estão sendo feitos na determinação dos mecanismos subjacentes e das proteínas regulatórias que tomam parte. Aqui, novamente, os estudos genéticos em leveduras foram fundamentais.

A estrutura da cromatina

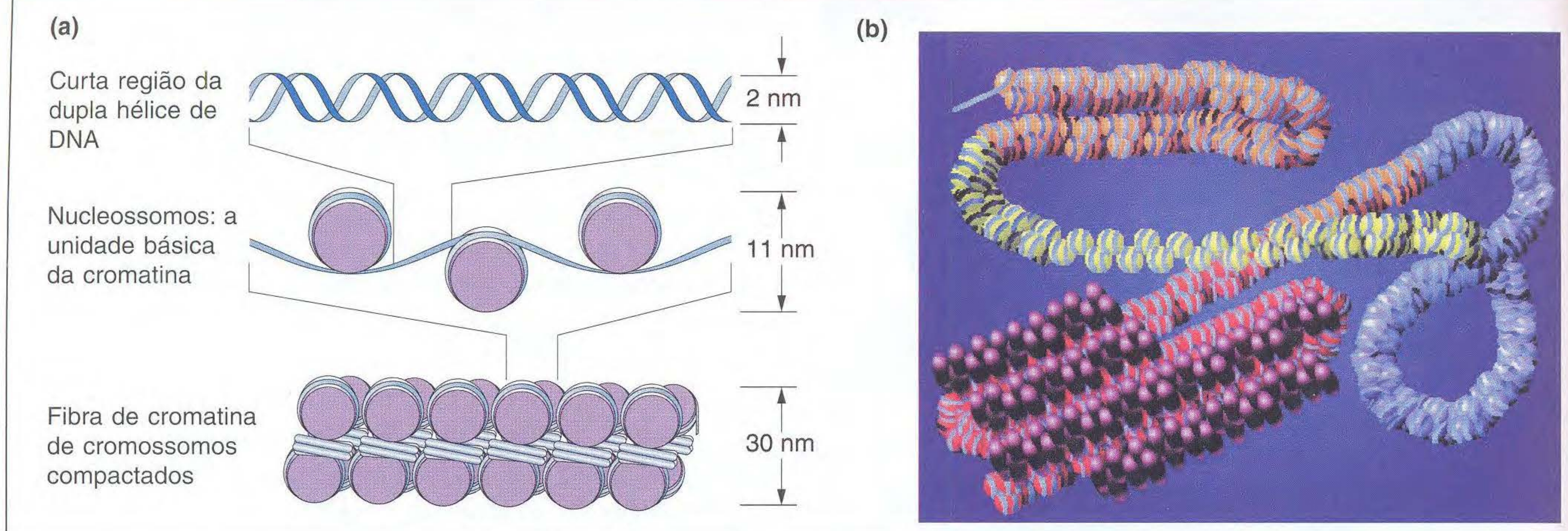


FIG. 11.10 (a) O nucleossomo na cromatina descondensada e condensada. (b) A estrutura da cromatina varia ao longo do tamanho de um cromossomo. A cromatina menos condensada (eucromatina) é mostrada em amarelo, as regiões de condensação intermediária estão em laranja e azul, e a heterocromatina revestida com proteínas especiais (púrpura) está em vermelho. [(b) De P. J. Horn and C. L. Peterson, "Chromatin Higher Order Folding: Wrapping Up Transcription," *Science* 297, 2002, 1827, Fig. 3. Copyright 2002, AAAS.]

Proteínas remodeladoras de cromatina e ativação gênica

Duas triagens genéticas em leveduras quanto a mutantes em processos aparentemente não-relacionados levaram à descoberta do mesmo gene cujo produto tem um papel central na remodelagem de cromatina. Em ambos os casos, as células de leveduras foram tratadas com agentes que causariam mutações. Em uma triagem, essas células de leveduras mutagenizadas foram triadas à busca de células que não podiam crescer bem em sucrose (mutantes não-fermentadores de sucrose, *snf*). Em outra triagem, as células de levedura mutagenizadas foram triadas quanto a mutantes que eram defectivos em mudar seu tipo reprodutivo (mutantes de mudança, *swi*; veja Seção 11.4). Muitos mutantes para loci diferentes foram recuperados em cada triagem, mas um gene mutante foi visto causando ambos os fenótipos. Os mutantes nos chamados locus *swi2/snf2* ("switch-sniff") não podiam usar efetivamente a sucrose nem o tipo reprodutivo.

Qual a ligação entre a habilidade em usar açúcar e a habilidade em mudar os tipos reprodutivos? A proteína Snf2-Swi2 foi purificada e descoberta como sendo parte de um grande complexo multissubunitário, chamado complexo SWI-SNF, que pode reposicionar os nucleossomos em uma análise de tubo de ensaio se o ATP for fornecido como fonte de energia (Figura 11.11). Em algumas situações, o complexo multissubunitário SWI-SNF ativa a transcrição movendo os nucleossomos que estão cobrindo as seqüências TATA e, desse modo, facilita a ligação da RNA polimerase II. O complexo SWI-SNF é, portanto, um coativador.

Gal4 também se liga ao complexo SWI-SNF e recruta o complexo de remodelagem de cromatina para promotores ativados. As linhagens de leveduras contendo um complexo defectivo SWI-SNF mostram um nível reduzido de atividade de Gal4. Por que um ativador deve usar vários mecanismos de ativação? Existem pelo menos dois motivos compreendidos

no momento. O primeiro é que a acessibilidade a promotores-alvo pode mudar em estágios diferentes do ciclo celular ou em tipos celulares diferentes (em eucariontes multicelulares). Por exemplo, durante a mitose, quando a cromatina é mais condensada, os genes são menos acessíveis. Nesse estágio, Gal4

O remodelamento da cromatina expõe seqüências reguladoras

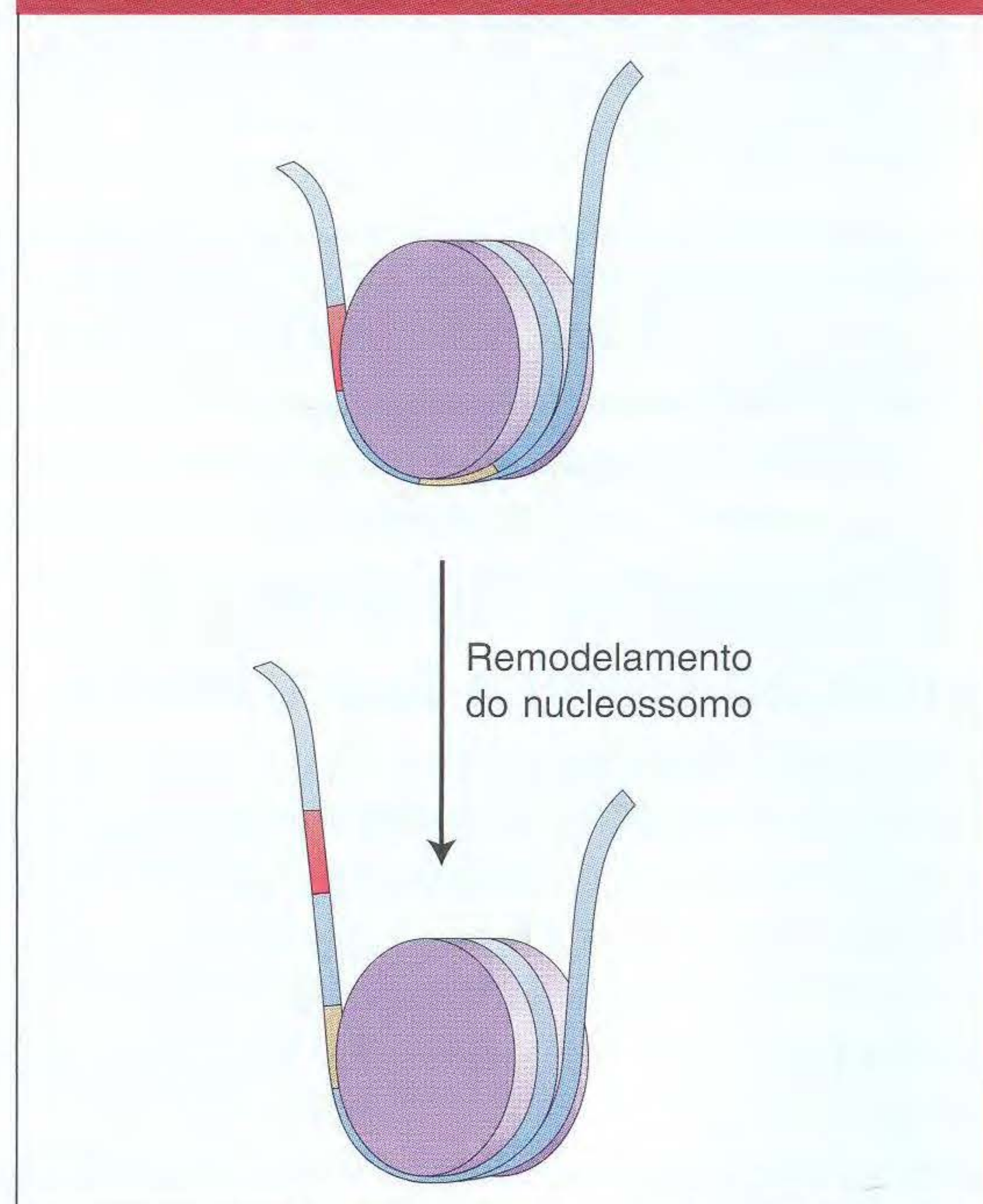


FIG. 11.11 O octâmero de histonas desliza em resposta à atividade de remodelamento da cromatina (tal como o complexo SWI-SNF), nesse caso expondo o DNA marcado em vermelho. (Veja a Figura 11.15 para detalhes de como SWI-SNF é recrutado para uma região particular do DNA.) [De J. Watson et al., *Molecular Biology of the Gene*, Quinta Edição, copyright © 2004, Benjamin Cummings.]

deve recrutar os complexos de remodelagem de cromatina, enquanto, em outros momentos, tal recrutamento pode não ser necessário para ativar a expressão gênica.

Um segundo motivo é que muitos fatores de transcrição atuam em combinações para controlar a expressão gênica sinergisticamente. Veremos resumidamente que essa sinergia combinatória é um resultado do fato de que os complexos de remodelagem de cromatina e a maquinaria transcrricional são recrutados mais eficientemente quando vários fatores de transcrição atuam juntos.

A cromatina pode ser dinâmica; os nucleossomos não estão em posições necessariamente fixas no cromossomo. A remodelagem de cromatina muda a densidade do nucleossomo ou a posição, e é uma parte integral da regulação gênica eucariótica.

Histonas e remodelagem de cromatina

Vamos observar mais atentamente o nucleossomo para ver se alguma parte dessa estrutura pode levar a informação

necessária para influenciar a posição do nucleossomo e/ou a densidade do nucleossomo.

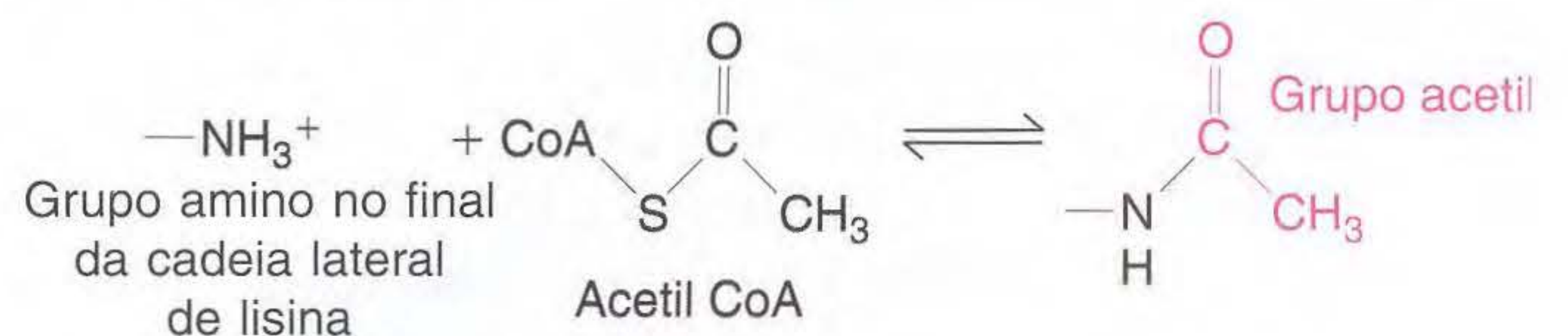
Um código de histona Como já foi visto, a maioria dos nucleossomos são compostos de um octâmero feito de duas cópias cada do cerne de quatro histonas. As histonas são conhecidas como sendo as proteínas mais conservadas na natureza; isto é, as histonas são quase idênticas em todos os organismos eucarióticos de leveduras a plantas e animais. Essa conservação contribuiu para uma visão de que as histonas não podem participar de nada mais complicado do que a embalagem do DNA para caber no núcleo. Entretanto, lembre que o DNA, com suas quatro bases, também foi considerado uma molécula muito “simples” para ser o mapa de todos os organismos do Planeta.

A Figura 11.12 mostra um modelo da estrutura do nucleossomo que representa as contribuições de muitos estudos. Devemos notar que as proteínas histonas estão organizadas no cerne do octâmero com suas pontas aminotermiais projetando-se do nucleossomo. Essas pontas que se projetam são chamadas de **caudas de histona**. Desde o início dos anos de 1960, lisinas específicas nas caudas de histonas foram conhecidas como sendo capazes de se modificar covalentemente pela ligação de grupos acetil e metil. Essas reações ocorrem após as proteínas histona terem sido traduzidas e mesmo após a histona ter sido incorporada ao nucleossomo.

Existem hoje em dia pelos menos 150 modificações diferentes de histonas que requerem uma ampla variedade de moléculas em adição aos grupos acetil e metil já mencionados (por exemplo, fosforilação e ubiquitinilação).

Acetilação, desacetilação de histonas e expressão gênica

A reação de acetilação é a melhor caracterizada das modificações das histonas:



Note que a reação é reversível, o que significa que os grupos acetil podem ser adicionados e removidos da mesma histona específica. Com 44 lisinas na histona disponíveis para aceitar o grupo acetil, a presença ou ausência desses grupos podem levar uma enorme quantidade de informação. Por esse motivo, a modificação covalente das caudas de histona é chamada de **código de histona**. Os cientistas criaram a expressão código de histona porque a modificação covalente das caudas de histona é remanescente do código genético. Para o código de histona, a informação é estocada nos padrões de modificação de histona e não na sequência de nucleotídeos. Com mais de 150 modificações de histona conhecidas, há um imenso número de padrões possíveis, e seus efeitos na estrutura da cromatina e regulação transcrricional estão apenas começando a ser decifrados. Para aumentar essa complexidade, o código provavelmente não é interpretado de modo preciso em todos os organismos. No momento, vamos ver

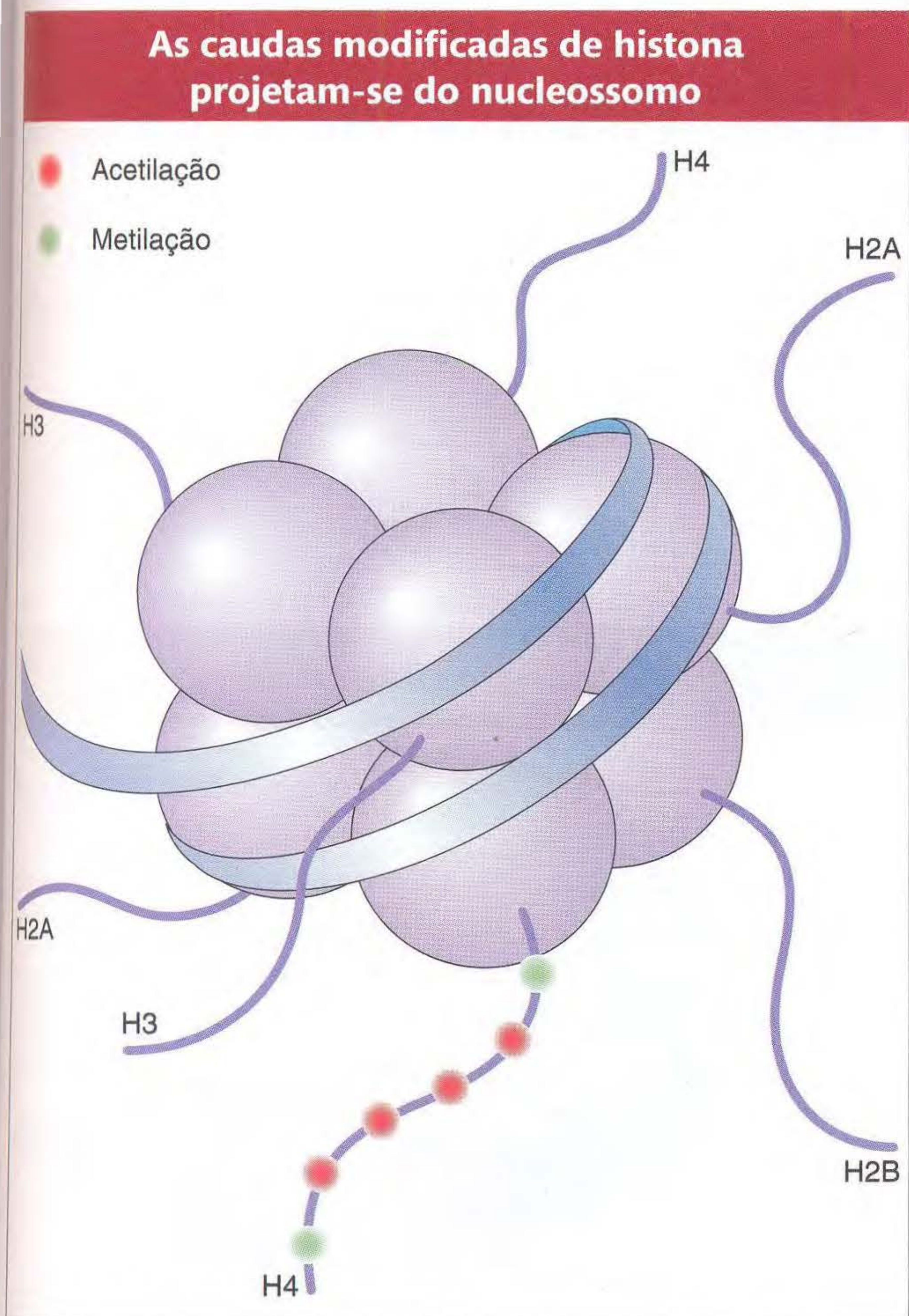


FIG. 11.12 Estrutura do nucleossomo mostrando sete das oito histonas e a maioria de suas caudas, mas não todas. Os sítios de modificações pós-traducionais, tais como a acetilação e a metilação, são mostrados para uma cauda de histona. De fato, todas as caudas contêm tais sítios.

como a acetilação de aminoácidos de histona influencia a estrutura e a expressão do gene.

Durante anos, têm se acumulado evidências de que as histonas associadas aos nucleossomos de genes ativos são ricas em grupos acetil (ditas **hiperacetiladas**), enquanto os genes inativos são subacetilados (**hipoacetilados**). A enzima responsável pela adição de grupos acetil, histona acetiltransferase (HAT), demonstrou-se muito difícil de isolar. Quando finalmente ela foi isolada e deduzida a sua sequência de proteína, foi descoberto que ela é um *ortólogo* de um ativador transcricional chamado GCN5 (significando que ela é codificada pelo mesmo gene em um organismo diferente). Assim, a conclusão foi de que GCN5 é uma histona acetiltransferase. Ela se liga ao DNA nas regiões reguladoras de alguns genes e ativa a transcrição acetilando histonas vizinhas. Vários complexos proteicos que são recrutados por ativadores transcacionais hoje são compreendidos como possuindo a atividade HAT.

Como a acetilação das histonas facilita mudanças na expressão gênica? Parece haver pelo menos dois mecanismos para fazer isso. Primeiro, a adição de grupos acetil a histonas específicas pode alterar a interação em um nucleossomo entre o DNA e um octâmero de histonas, de modo que o octâmero mais provavelmente desliza ao longo do DNA para uma nova posição. Segundo, a acetilação de histonas, em conjunto com outras modificações de histona, influencia a ligação de proteínas reguladoras ao DNA. A proteína reguladora ligada pode tomar parte de várias funções que, direta ou indiretamente, aumentam a frequência de início da transcrição.

Como outras modificações de histona, a acetilação é reversível, e as **histonas desacetilase (HDAT)** também foram identificadas. Tais proteínas têm papéis importantes na repressão gênica. Por exemplo, na presença de galactose ou glicose, a ativação de genes *GAL* é evitada pela proteína Mig1. Mig1 é um repressor de ligação a sequências específicas de DNA que se ligam a um sítio entre o elemento UAS e o promotor do gene *GAL1* (Figura 11.13). Mig1 recruta um complexo proteico chamado Tup1 que contém uma histona desacetilase e que reprime a transcrição gênica. O complexo

Tup1 é um exemplo de um **co-repressor** que facilita a expressão gênica mas, em si, não é um repressor de ligação ao DNA. O complexo Tup1 também é recrutado por outros repressores de leveduras, tais como MAT α 2 (veja adiante), e contrapartes desse complexo são encontradas em todos os eucariontes.

Na maioria dos casos examinados, a acetilação e desacetilação de histonas promovem e reprimem a transcrição gênica, respectivamente. Essas atividades são recrutadas para genes por ativadores específicos de sequência e repressores.

► 11.4 Mecanismo de Ação Acentuadora

O desenvolvimento de um organismo complexo requer que os níveis de transcrição sejam regulados em uma ampla gama. Pense no mecanismo de regulação como sendo um reostato em vez de um interruptor. Nos eucariontes, os níveis de transcrição são finamente ajustáveis agrupando os sítios de ligação em acentuadores. Vários fatores de transcrição diferentes ou várias moléculas do mesmo fator de transcrição podem se ligar a sítios adjacentes. A ligação desses fatores a sítios que estão separados pela distância correta leva a um efeito amplificado, ou superaditivo, em ativar a transcrição. Quando um efeito é maior que o aditivo, ele é chamado de **sinérgico**.

A ligação de várias proteínas reguladoras a múltiplos sítios de ligação em um acentuador pode catalisar a formação de um **acentuassomo** (*enhanceosome*), um grande complexo de proteínas que pode agir sinérgicamente para ativar a transcrição. Na Figura 11.14, você pode ver como proteínas arquitetonas dobram o DNA para promover interações cooperativas entre outras proteínas de ligação ao DNA. Nesse modo de ação do acentuassomo, a transcrição é

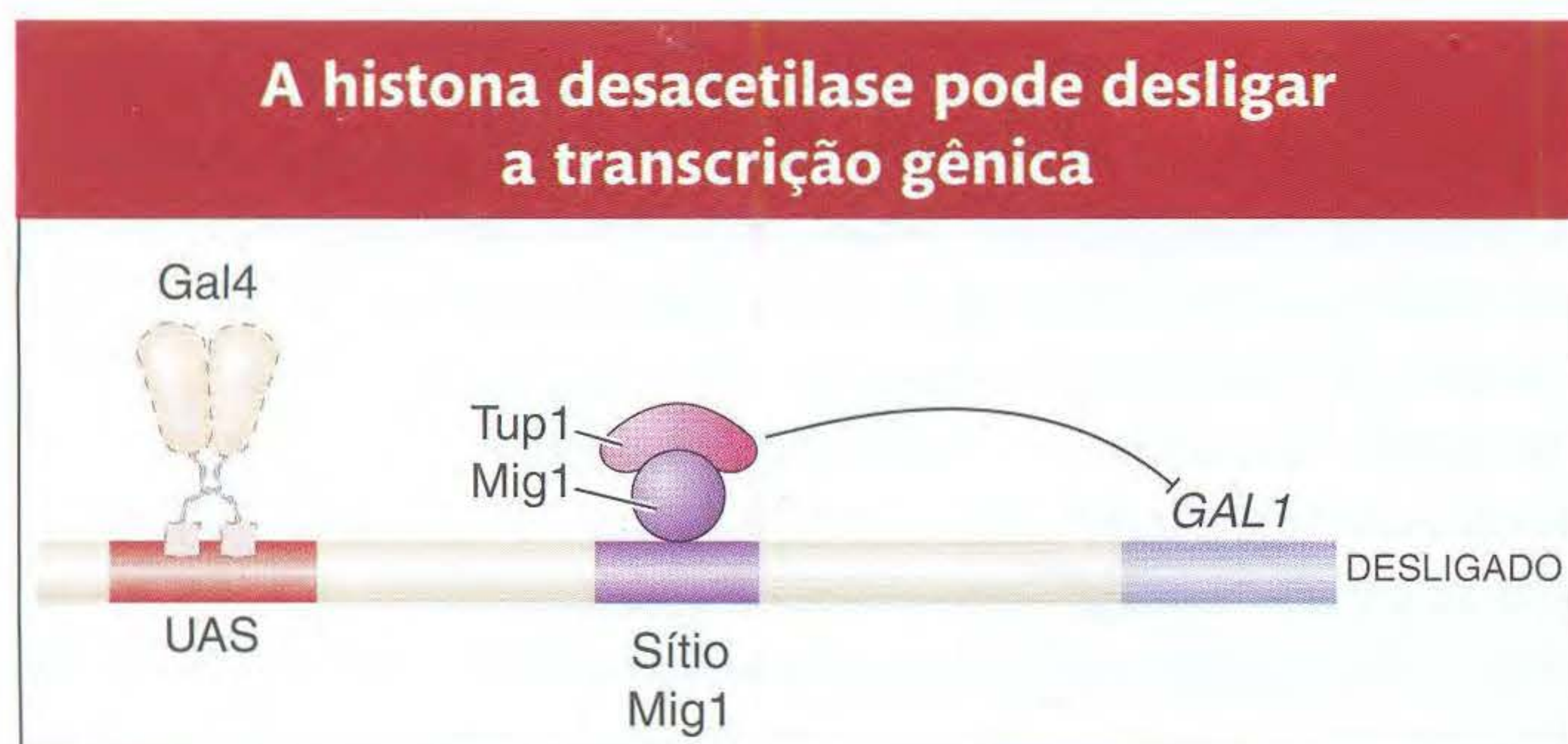


FIG. 11.13 O recrutamento de um complexo de repressão leva à repressão da transcrição. Na presença de glicose, a transcrição de *GAL1* é reprimida pela proteína Mig1, que se liga a um sítio entre UAS e o promotor do gene *GAL1*. Mig1 recruta o complexo de repressão Tup1, o qual recruta uma histona desacetilase, desligando a transcrição gênica. [De J. Watson et al., *Molecular Biology of the Gene*, Quinta Edição, copyright © 2004, Benjamin Cummings.]

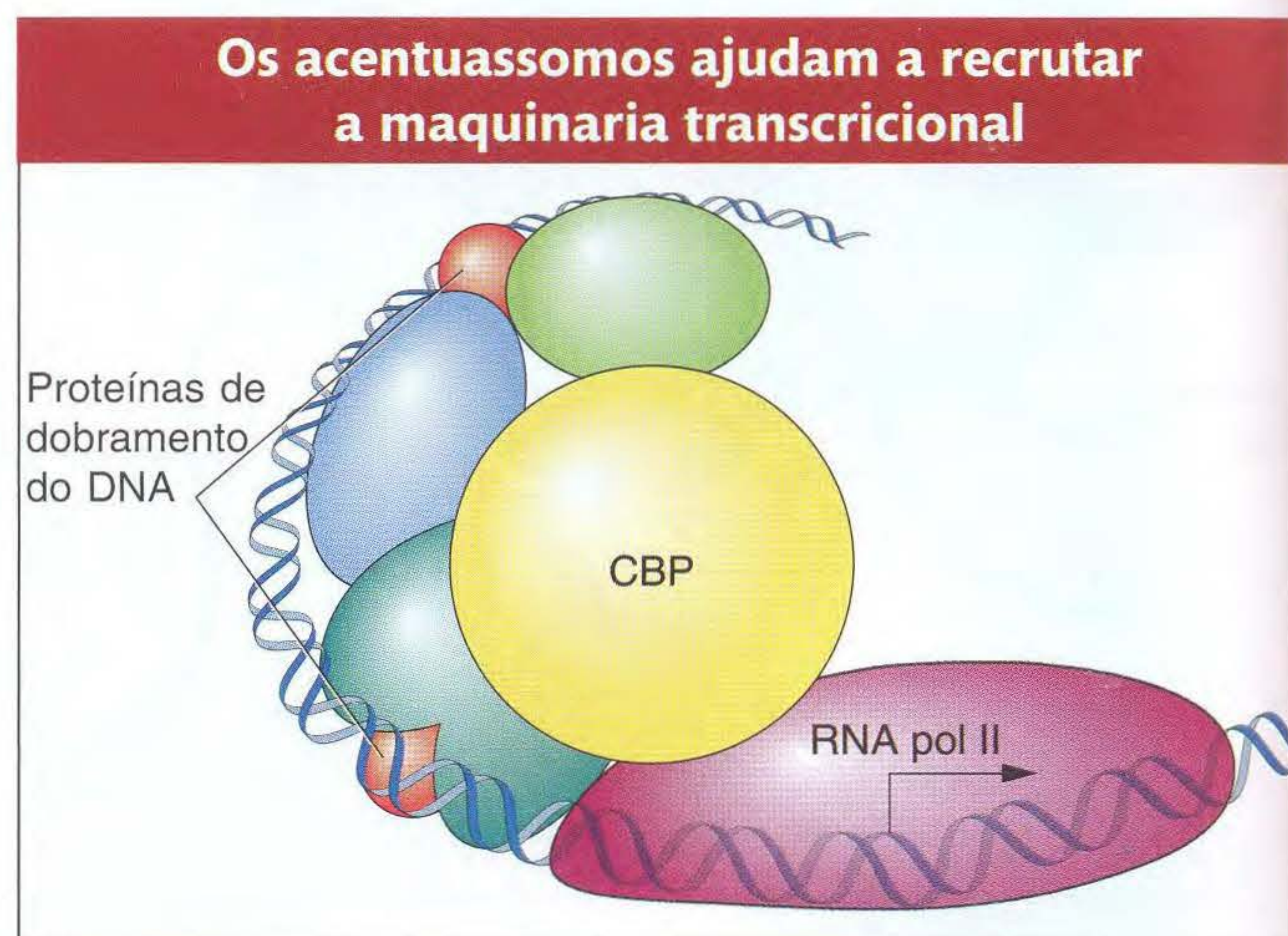


FIG. 11.14 O acentuassomo β -interferon. Nesse caso, os fatores de transcrição recrutam um coativador (CBP), que se liga tanto aos fatores de transcrição quanto à RNA polimerase II, iniciando a transcrição. [De A. J. Courey, "Cooperativity in Transcriptional Control," *Curr. Biol.* 7, 2001, R250-R253, Fig. 1.]

ativada a níveis muito altos apenas quando todas as proteínas estão presentes e tocando uma a outra do modo certo.

Para compreendermos melhor o que um acentuassomo é e como ele atua sinergisticamente, vejamos um exemplo específico.

O acentuassomo β -interferon

O gene humano de β -interferon, que codifica a proteína antiviral interferon, é um dos genes mais bem caracterizados em eucariontes. Ele normalmente está desligado, mas é ativado a níveis muito altos de transcrição na infecção viral. A chave para a ativação desse gene é a montagem de fatores de transcrição em um acentuassomo acerca de 100 pb do TATA box e sítio de início da transcrição. As proteínas regulatórias do acentuassomo β -interferon ligam-se à mesma face da dupla hélice de DNA. Ligadas ao outro lado da hélice estão várias proteínas estruturais que dobram o DNA e permitem que proteínas reguladoras diferentes se toquem e formem um complexo ativado. Quando todas as proteínas reguladoras estão ligadas e interagindo corretamente, elas formam uma “plataforma”, um sítio de ligação de alta afinidade para a proteína CBP, uma proteína coativadora que também recruta a maquinaria de transcrição. A grande proteína CBP também contém uma atividade intrínseca de histona acetilase que modifica os nucleossomos e facilita altos níveis de transcrição.

Embora o promotor do β -interferon seja mostrado sem os nucleossomos na Figura 11.14, o acentuassomo é circundado por dois nucleossomos, chamados nuc 1 e nuc 2 na Figura 11.15. Um deles, nuc 2, está estrategicamente posicionado no TATA box e sítio de início da transcrição. Entretanto, a ligação de GCN5, outro coativador, é agora conhecida como precedendo a ligação de CBP. GCN5 acetila os dois nucleossomos. Após a acetilação, os fatores ativadores de transcrição recrutam o coativador CBP, a holoenzima RNA pol II e o complexo de remodelamento de cromatina SWI-SNF. SWI-SNF é então posicionado para afastar o nucleossomo 37 pb do TATA box, tornando o TATA box acessível à proteína de ligação TATA e permitindo que a transcrição seja iniciada.

As interações cooperativas ajudam a explicar várias observações curiosas sobre os acentuadores. Por exemplo, elas explicam por que, mudando qualquer fator de transcrição ou sítio de ligação, a atividade do acentuador é marcadamente reduzida. Elas também explicam por que a distância entre os sítios de ligação dentro do acentuador é uma característica tão crítica. Além disso, os acentuadores não têm que estar perto do sítio de início da transcrição, como no exemplo mostrado na Figura 11.15. Uma característica dos acentuadores é que eles podem ativar a transcrição quando estão situados a grandes distâncias do promotor (>50 kb), seja antecedentes ou posteriores a um gene ou mesmo em um íntron.

Os acentuassomos recrutam remodeladores de cromatina

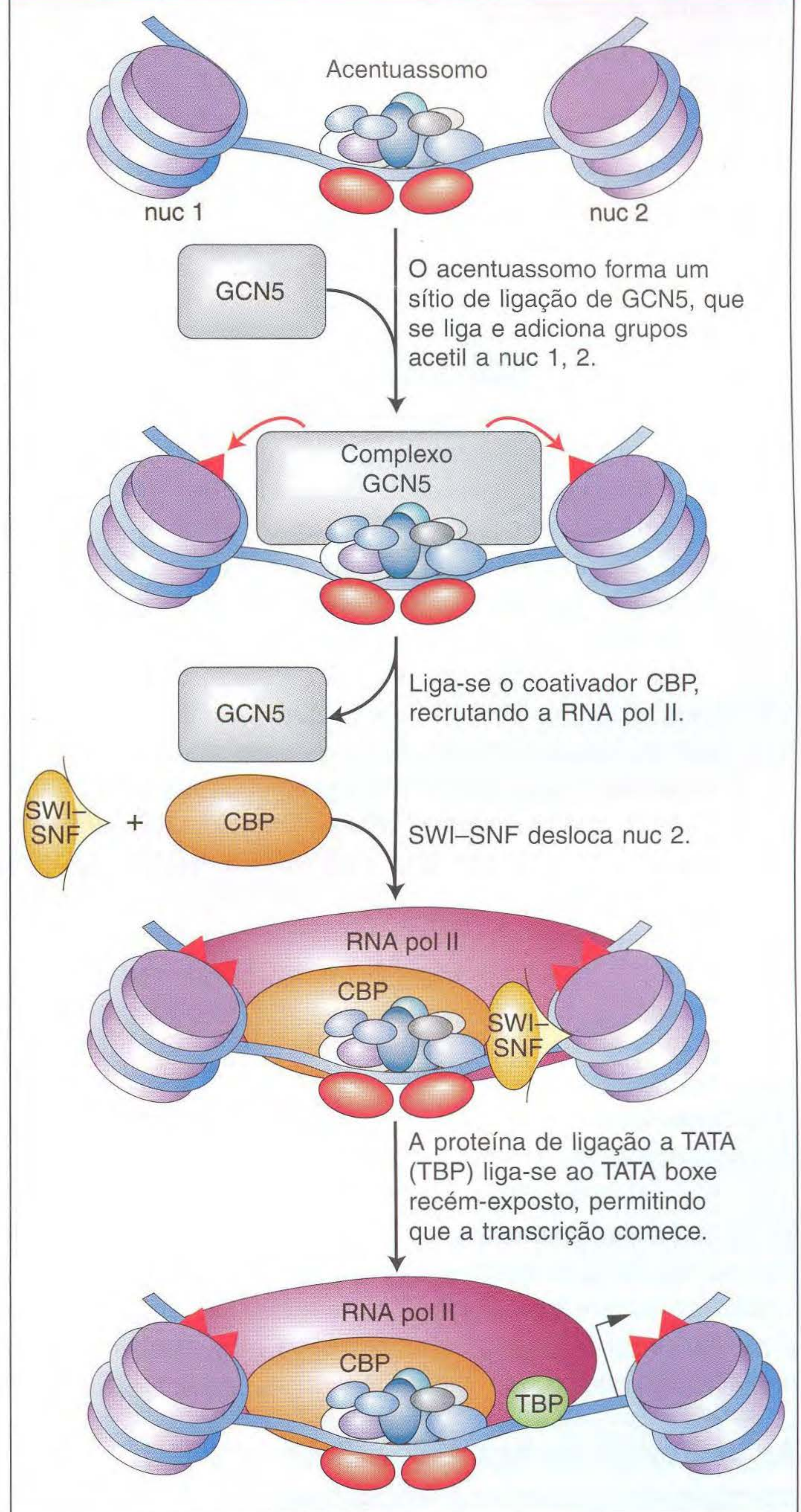


FIG. 11.15 O acentuassomo β -interferon atua para mover os nucleossomos recrutando o complexo SWI-SNF.

Os acentuadores eucarióticos podem agir a grandes distâncias para modular a atividade do aparato de transcrição. Os acentuadores contêm sítios de ligação para muitos fatores de transcrição, que se ligam e interagem cooperativamente. Essas interações resultam em uma variedade de respostas, incluindo o recrutamento de coativadores adicionais e remodelagem da cromatina.

O controle do tipo reprodutivo de levedura: Interações combinatórias

Até agora, enfocamos neste capítulo a regulação de genes isolados ou alguns genes em uma via. Em organismos multicelulares, tipos celulares distintos diferem na expressão de centenas de genes. A expressão ou repressão de grupos de genes deve, portanto, ser coordenada na produção de tipos particulares de células. Um dos exemplos mais bem conhecidos de regulação do tipo de célula em eucariontes é a regulação do tipo reprodutivo em leveduras. Esse sistema regulador foi dissecado por uma elegante combinação de genética, biologia molecular e bioquímica. O tipo reprodutivo serve como um excelente modelo para compreender a lógica da regulação gênica em animais multicelulares.

A levedura *Saccharomyces cerevisiae* pode existir em qualquer dos três tipos diferentes de células conhecidos como **a**, **α** e **a/α** (veja o Capítulo 2). Os dois tipos celulares **a** e **α** são haplóides e contêm uma cópia de cada cromossomo. Embora as duas células haplóides não possam ser distintas por seu aspecto ao microscópio, elas podem ser diferenciadas por várias características celulares específicas, principalmente seu tipo reprodutor (veja o box Organismo-modelo *Leveduras*). Uma célula **α** só se reproduz com uma célula **a** e secreta um **feromônio** oligopeptídico, ou hormônio sexual, chamado de fator **α**, que bloqueia as células **a** no ciclo celular. Uma célula do tipo **a** só se reproduz com uma célula **α** e secreta um feromônio, chamado de fator **a**, que bloqueia as células **α**. A célula diplóide **a/α** não se reproduz, é maior do que as células **α** e **a** e não responde a hormônios reprodutivos.

A análise genética de mutantes defeituosos em reprodução mostrou que o tipo celular é controlado por um único locus genético, o locus do tipo reprodutor, *MAT*. Existem dois alelos do locus *MAT*: as células **a** haplóides têm o alelo *MATa*, as células haplóides **α** têm o alelo *MATα* e a diplóide **a/α** tem ambos os alelos. Embora o tipo reprodutivo esteja sob controle genético, algumas linhagens mudam seu tipo reprodutor, às vezes tão freqüentemente quanto cada divisão celular. Examinaremos a base da mudança mais adiante neste capítulo, mas, primeiro, vejamos como cada tipo de célula expressa o conjunto correto de genes.

As proteínas de ligação ao DNA regulam combinatoriamente a expressão de genes específicos do tipo de célula

Como o locus *MAT* controla o tipo de célula? A análise genética dos mutantes que não podem reproduzir-se identificaram vários genes estruturais que são separados do locus *MAT*, mas seus produtos proteicos são necessários para a reprodução. Um grupo de genes estruturais só é expresso no tipo celular **α** (genes **α**-específicos), e outro grupo só é expresso no tipo celular **a** (genes **a**-específicos). Os alelos diferentes do locus *MAT* codificam proteínas reguladoras que controlam qual desses conjuntos de genes estruturais é expresso em cada tipo de célula. Em adição, uma proteína reguladora não codificada pelo locus *MAT*, chamada MCM1, tem um papel principal em regular o tipo de célula.

O caso mais simples é o tipo celular **a** (Figura 11.16a). O locus *MATa* codifica uma única proteína reguladora, **a1**. Entretanto, essa proteína reguladora não tem efeito nas células haplóides, mas apenas nas células diplóides. Em uma célula haplóide **a**, a proteína reguladora MCM1 liga a expressão dos genes estruturais necessários a uma célula **a**, ligando-se com seqüências reguladoras dentro de promotores gênicos específicos de **a**.

Em uma célula **α**, os genes estruturais **α**-específicos devem ser transcritos, mas, em adição, a proteína MCM1 deve ser impedida de ativar os genes **a**-específicos. A seqüência de DNA do alelo *MATα* codifica duas proteínas, **α1** e **α2**, que são produzidas por unidades de transcrição separadas. Essas duas proteínas têm papéis reguladores diferentes na célula **α**, como pode ser demonstrado analisando suas propriedades de ligação ao DNA *in vitro* (Figura 11.16b). A proteína **α1** liga-se, em conjunto com a proteína MCM1, a uma pequena seqüência de DNA que controla vários genes **α**-específicos. Assim, **α1** é um ativador da expressão gênica **α**-específica. A proteína **α2** reprime a transcrição de genes **a**-específicos. Ela se liga a um dímero, com MCM1, a sítios em seqüências de DNA situadas a 5' de um grupo de genes **a**-específicos e atua como um repressor.

Em uma célula diplóide de levedura, todas as três proteínas reguladoras codificadas pelo locus *MAT* são expressas (Figura 11.16c). Qual é o resultado? A proteína **a1** codificada por *MATa* tem um papel a cumprir finalmente. A proteína **a1** pode se ligar a **α2** e alterar sua especificidade de ligação tal que o complexo **a1-α2** não se ligue a genes **a**-específicos. O complexo **a1-α2** liga-se a uma seqüência diferente encontrada antecedente a outro conjunto de genes, chamados de específicos haplóides, que são expressos em células haplóides, mas não em células diplóides. Nas células diplóides, então, **α2** existe em duas formas: (1) como um complexo **α2-MCM1** que reprime genes **a**-específicos e (2) em um complexo com **a1** que reprime genes específicos haplóides. Os diferentes parceiros de ligação determinam quais seqüências específicas de DNA estão ligadas e quais genes são regulados por cada complexo contendo **α2**. A regulação de conjuntos diferentes de genes-alvo pela associação do mesmo fator de transcrição a parceiros diferentes de ligação tem um papel importante na geração de padrões diferentes de expressão gênica em tipos de células diferentes dentro de eucariontes multicelulares.

Em leveduras e em eucariontes multicelulares, os padrões específicos do tipo de célula da expressão gênica são controlados por combinações de fatores de transcrição que interagem.

Isoladores de bloqueio de acentuador

Um elemento regulatório, tal como um acentuador, que pode agir em dezenas de milhares de pares de bases, pode interferir na regulação de genes vizinhos. Para evitar tal ativação promíscua, desenvolveram-se elementos reguladores chamados de **isoladores de bloqueio de acentuador**. Quando posicionados

As combinações de proteínas reguladoras controlam os tipos celulares

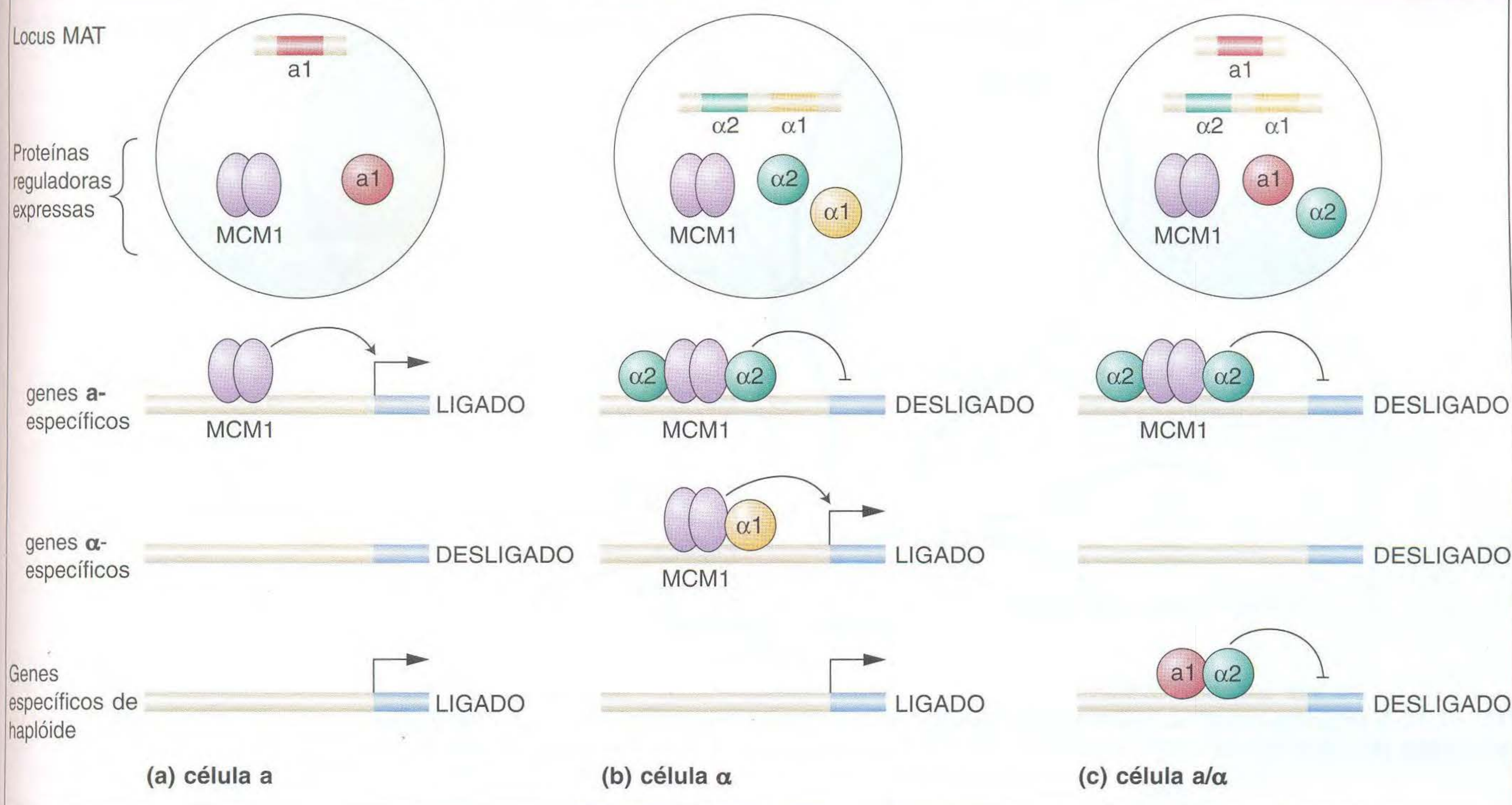


FIG. 11.16 Controle da expressão gênica específica do tipo de célula em levedura. Os três tipos de *S. cerevisiae* são determinados pelas proteínas reguladoras a1, α1 e α2, que regulam diferentes subgrupos de genes-alvo. A proteína MCM1 atua em todos os três tipos de células e interage com α1 e α2.

entre um acentuador e um promotor, os isoladores de bloqueio de acentuador impedem que o acentuador ative a transcrição desse promotor. Tais isoladores não têm efeito na ativação de outros promotores que não estão separados de seus acentuadores pelo isolador (Figura 11.17). Vários modelos foram propostos para explicar como um isolador pode bloquear a atividade de um acentuador apenas quando colocado entre um acentuador e um promotor. Muitos dos modelos, como o mostrado na Figura 11.18, propõem que o DNA é organizado em alças contendo genes ativos. De acordo com esse modelo, os isoladores atuam movendo um promotor para uma nova alça, onde é protegido do acentuador.

Como você verá em seguida, os isoladores de bloqueio de acentuador são um componente fundamental de um fenômeno chamado de imprinting genômico.

► 11.5 Imprinting Genômico

O fenômeno de **imprinting genômico** foi descoberto há quase 20 anos em mamíferos. No imprinting genômico, alguns genes autossômicos têm padrões incomuns de herança. Por exemplo, um alelo *igf2* é expresso em um camundongo apenas se for herdado do pai do camundongo, um exemplo de **imprinting materno** porque uma cópia do gene derivada da mãe é inativa. Contrariamente, um alelo *H19* de camundongo só é expresso se for herdado da mãe. *H19* é um exemplo de **imprinting paterno** porque a cópia paterna é inativa. A consequência do imprinting parental é que os genes imprintados são expressos como se fossem a única cópia do gene presente na célula, muito embora existam duas. Assim, o imprinting é

Os isoladores de bloqueio de acentuador impedem ativação de acentuador

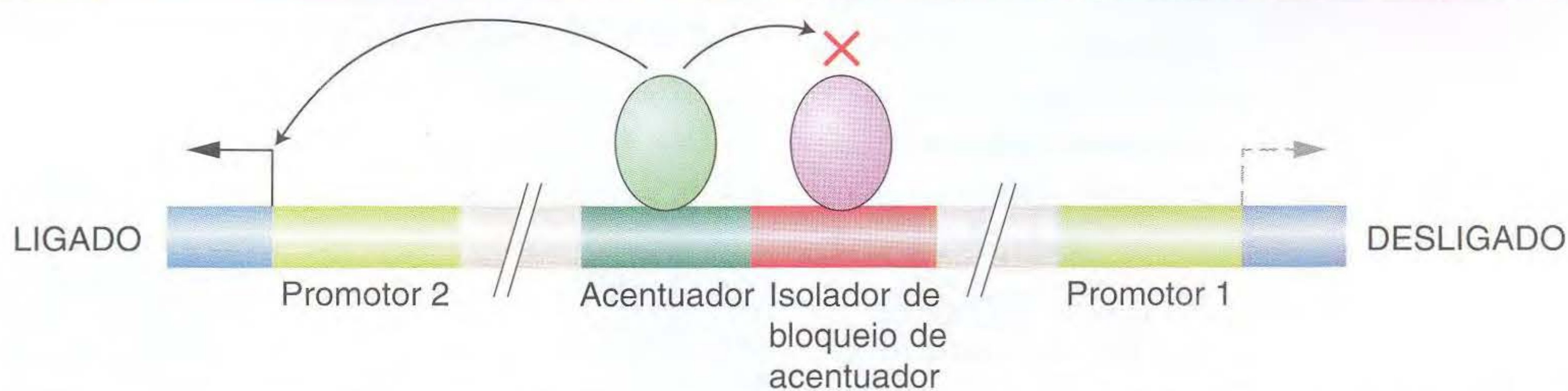


FIG. 11.17 Os isoladores de bloqueio de acentuador impedem a ativação gênica quando colocados entre um acentuador e um promotor. [De M. Gaszner and G. Felsenfeld, "Insulators: Exploiting Transcriptional and Epigenetic Mechanisms," *Nat. Rev. Genet.* 7, 2006, 703–713.]

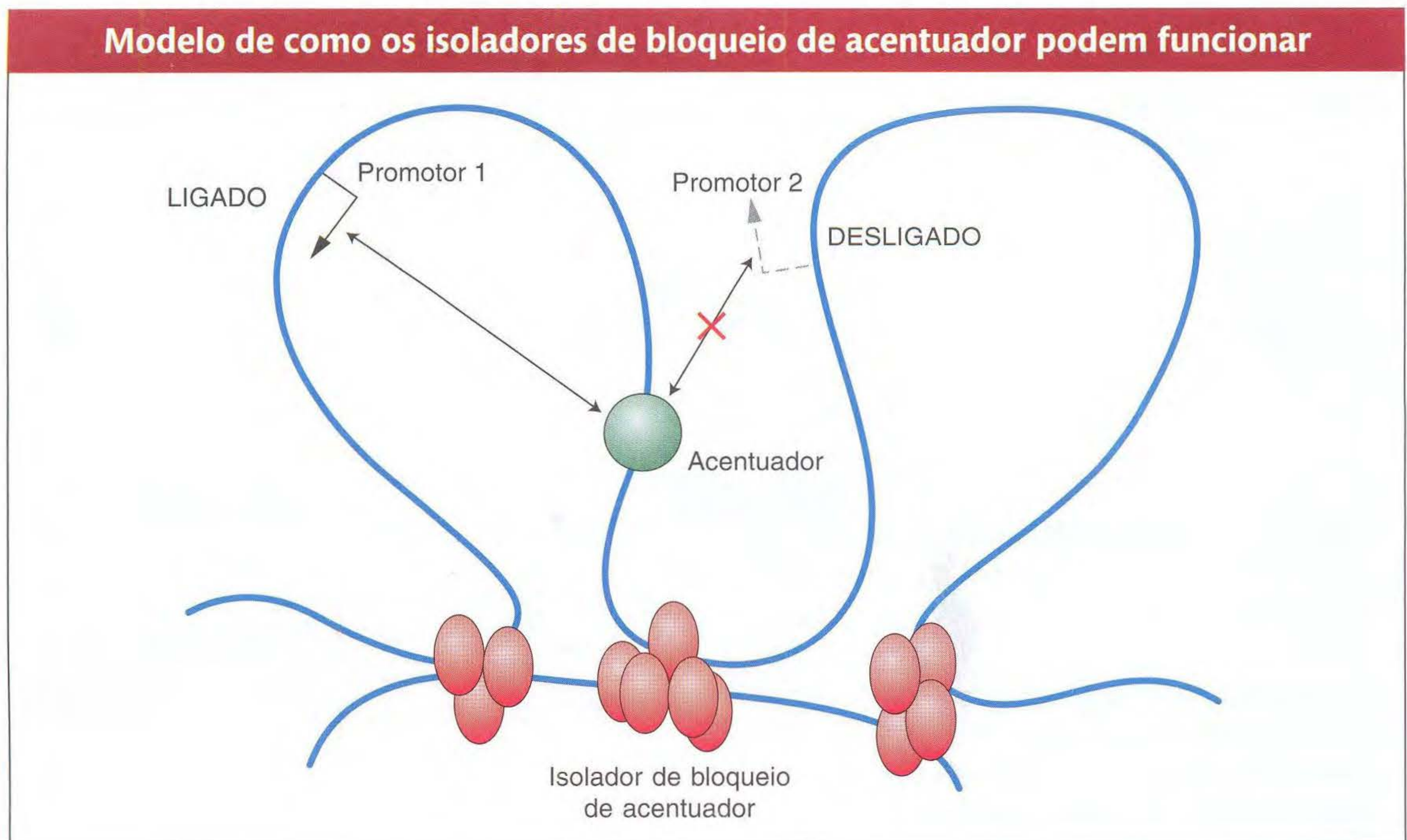
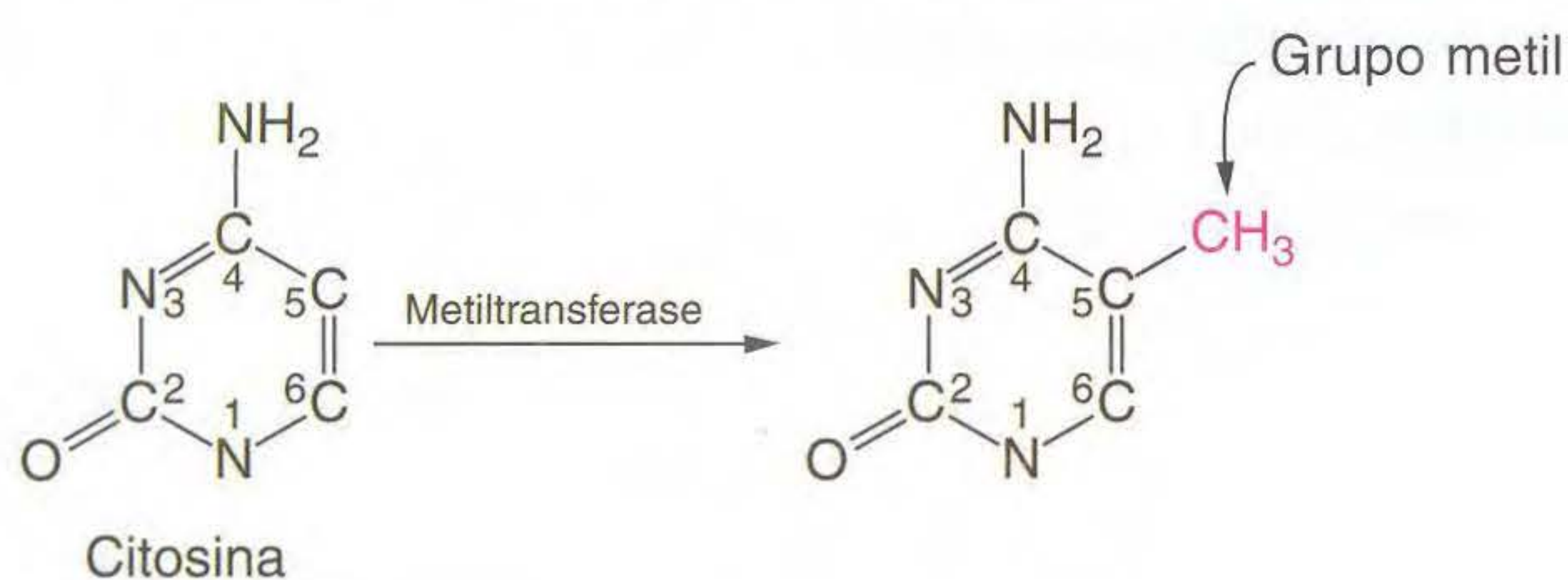


FIG. 11.18 A proposta é que os isoladores de bloqueio de acentuador (EB) criem novas alças que separam fisicamente um promotor de seu acentuador (E). [De M. Gaszner and G. Felsenfeld, "Insulators: Exploiting Transcriptional and Epigenetic Mechanisms," *Nat. Rev. Genet.* 7, 2006, 703–713.]

um exemplo de **herança monoalélica**. Importante, nenhuma mudança é observada nas seqüências de DNA de genes imprintados; isto é, o gene idêntico pode ser ativo ou inativo na prole, conforme tenha sido herdado da mãe ou do pai.

Se a seqüência de DNA do gene não está correlacionada à atividade, o que ocorre? A resposta é que o DNA nas regiões reguladoras de genes imprintados é metilada de modo específico do sexo no desenvolvimento de gametas. A **metilação do DNA** em geral resulta da adição enzimática de grupos metil à posição do carbono 5 de uma citosina específica.



Tanto as marcas de metilação do DNA quanto as marcas de modificação de histonas podem ser estavelmente herdadas de uma geração celular para a seguinte. Veremos depois, neste capítulo, como tais marcas são consideradas duplicadas no curso da replicação do DNA. Por enquanto, é suficiente dizer que tal alteração herdável, na qual a própria seqüência de DNA é inalterada, é chamada de **herança epigenética**, e as alterações (incluindo tanto a metilação do DNA quanto as modificações de histonas) são chamadas de **marcas epigenéticas**.

Voltemos novamente aos genes *ifg2* e *H19* de camundongo para ver como funciona o imprinting ao nível molecular. Esses dois genes estão situados em um grupo de genes imprintados no cromossomo 7 de camundongo. Existe uma

estimativa de 100 genes imprintados no camundongo, e a maioria é encontrada em grupos que compreendem de 3 a 11 genes imprintados. (Os humanos têm a maioria dos mesmos genes imprintados que os de camundongos.) Em todos os casos examinados, há um padrão específico de metilação de DNA para cada cópia de um gene imprintado. Para o grupo *ifg2-H19*, uma região específica do DNA que fica entre os dois genes (Figura 11.19) é metilada nas células germinativas masculinas e não-metilada nas células germinativas femininas. Essa região é chamada de região de controle de imprinting (ICR). Apenas a ICR não-metilada (feminina) pode ligar-se a uma proteína regulatória chamada CTCF. Quando ligada, a CTCF atua como um isolador de bloqueio de acentuador que impede a ativação do acentuador da transcrição de *Igf2*. Entretanto, o acentuador nas fêmeas ainda pode ativar a transcrição de *H19*. Nos machos, a CTCF não pode ligar-se à ICR, e o acentuador pode ativar a transcrição de *Igf2* (lembre que os acentuadores podem agir a grandes distâncias). O acentuador, entretanto, não pode ativar *H19*, pois a região metilada estende-se para o promotor *H19*. O promotor metilado não pode ligar-se às proteínas necessárias para a transcrição de *H19*.

Assim, vemos como um isolador de bloqueio de acentuador (nesse caso, CTCF ligada a parte de ICR) impede o acentuador de ativar um gene distante (nesse caso, *Igf2*). Além disso, vemos que o sítio de ligação de CTCF é metilado apenas nos cromossomos derivados do genitor masculino. A metilação do sítio de ligação de CTCF impede a ligação de CTCF nos machos e permite que o acentuador ative *Igf2*.

Note que o imprinting parental pode afetar muito a análise de heredogramas. Como o alelo herdado de um genitor é inativo, uma mutação no alelo herdado de um genitor parecerá

O imprinting genômico requer isoladores

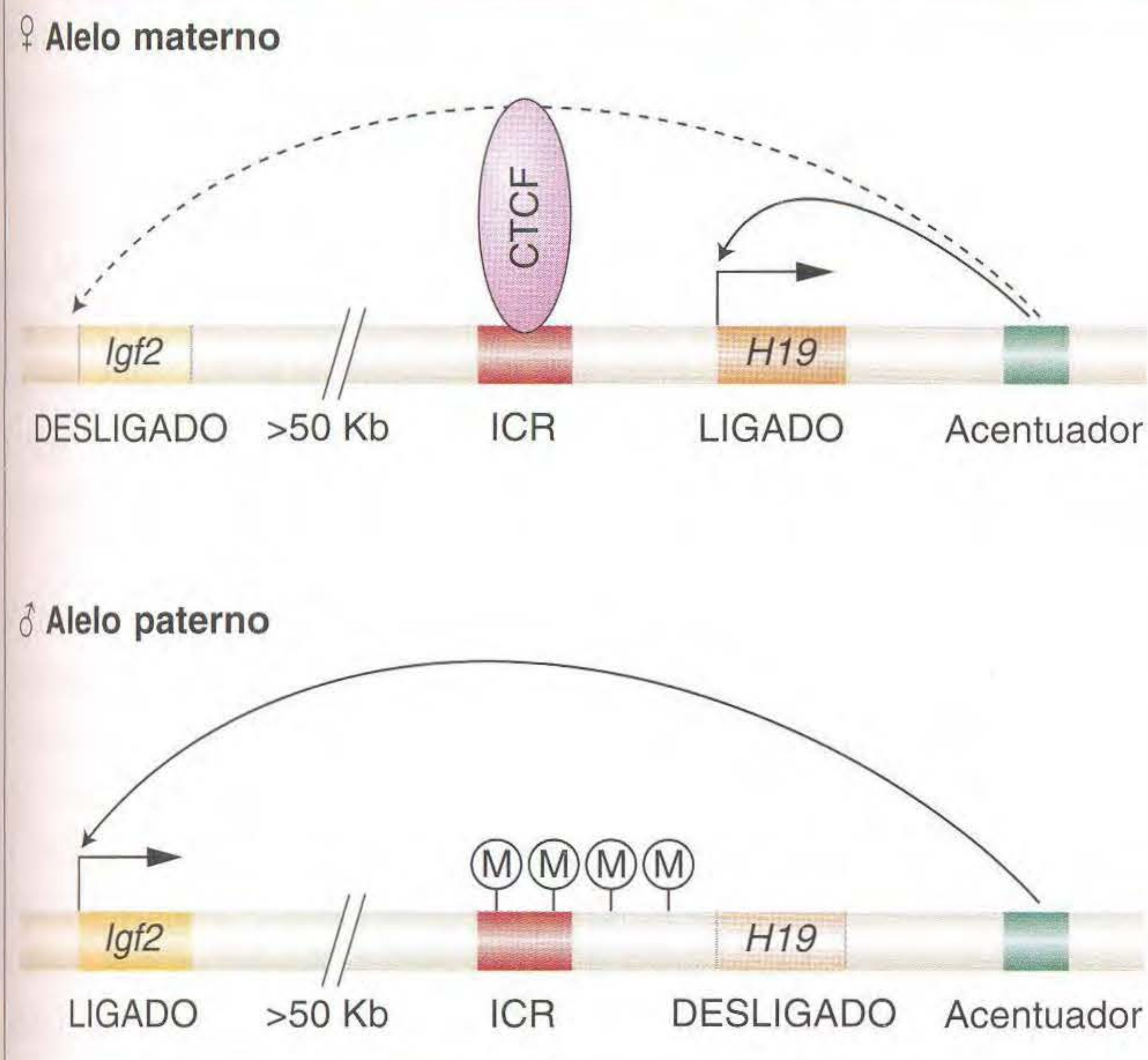


FIG. 11.19 Imprinting genômico no camundongo. A região de controle do imprinting (ICR) não está metilada nos gametas femininos e pode se ligar a um dímero CTCF, formando um isolador que bloqueia a ativação do acentuador de *Igf2*. A metilação (M) de ICR nas células germinativas masculinas impede a ligação CTCF, mas também impede a ligação de outras proteínas ao promotor *H19*.

ser dominante, enquanto, de fato, o alelo é expresso porque apenas um dos dois homólogos é ativo para esse gene. A Figura 11.20 mostra como uma mutação em um gene imprintado pode ter resultados diferentes no fenótipo do organismo se herdado do genitor masculino ou do feminino.

Muitas etapas são necessárias para o imprinting (Figura 11.21). Logo após a fertilização, os mamíferos separam células que irão tornar-se células germinativas. Os imprints são removidos ou apagados antes que as células germinativas sejam formadas. Sem sua marca distinta de metilação do DNA, esses genes agora são ditos *epigeneticamente equivalentes*. À medida que essas células germinativas primordiais se tornam gametas totalmente formados, os genes recessivos imprintados recebem a marca específica do sexo que irá determinar se o gene será ativo ou silencioso após a fertilização.

Mas, e quanto a Dolly e outros mamíferos clonados?

O imprinting genômico leva ao que muitos pensavam ser um requisito para a participação das células germinativas masculina e feminina no desenvolvimento de embriões de mamíferos. Isto é, os gametas masculino e feminino contêm subgrupos diferentes de genes impressos, de modo que o embrião terá um complemento completo de genes ativos imprintados. Por que, então, mamíferos tais como Dolly e, mais recentemente, porcos clonados, gatos, cães e vacas que foram derivadas de núcleos somáticos são capazes de sobreviver e desenvolver-se? Como já foi notado, a mutação mesmo

Herança incomum de genes imprintados

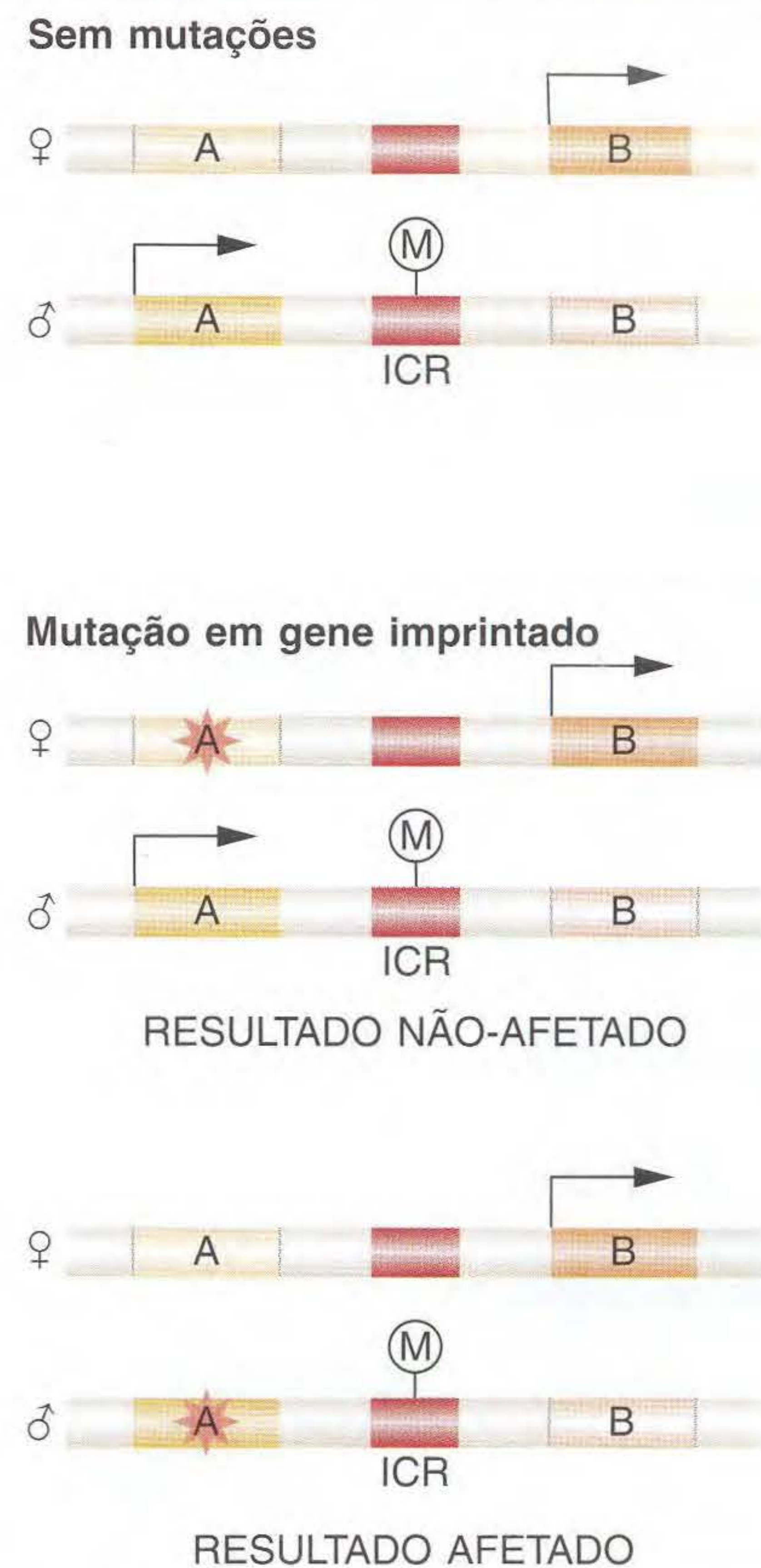


FIG. 11.20 Uma mutação (representada pela estrela laranja) no gene A não terá efeito se herdada do macho. Abreviações: M, metilação; ICR, região de controle do imprinting. [De S. T. da Rocha and A. C. Ferguson-Smith, "Genomic Imprinting," *Curr. Biol.* 14, 2004, R646-R649.]

de um único gene imprintado pode ser letal ou pode levar a uma doença grave.

No momento, os cientistas não compreendem por que a clonagem de muitas espécies de mamíferos tem tido sucesso. Entretanto, a despeito desses sucessos, a clonagem é extremamente ineficiente em todas as espécies testadas. Para a maioria dos experimentos, um clone bem-sucedido é um evento extremamente raro, requerendo centenas ou, mesmo, milhares de tentativas. Podemos pensar que a falha da maioria dos embriões clonados em se desenvolver em organismos viáveis é um atestado da importância dos mecanismos epigenéticos de regulação gênica nos eucariontes. Assim, isso ilustra como o conhecimento da sequência completa do DNA de todos os genes em um organismo é apenas uma primeira etapa na compreensão de como os genes eucarióticos são regulados.

11.6 Domínios de Cromatina e Sua Herança

Até agora, vimos como os genes são ativados em um ambiente de cromatina. Entretanto, como visto no início deste capítulo, a maioria dos genes em genomas eucarióticos estão desliga-

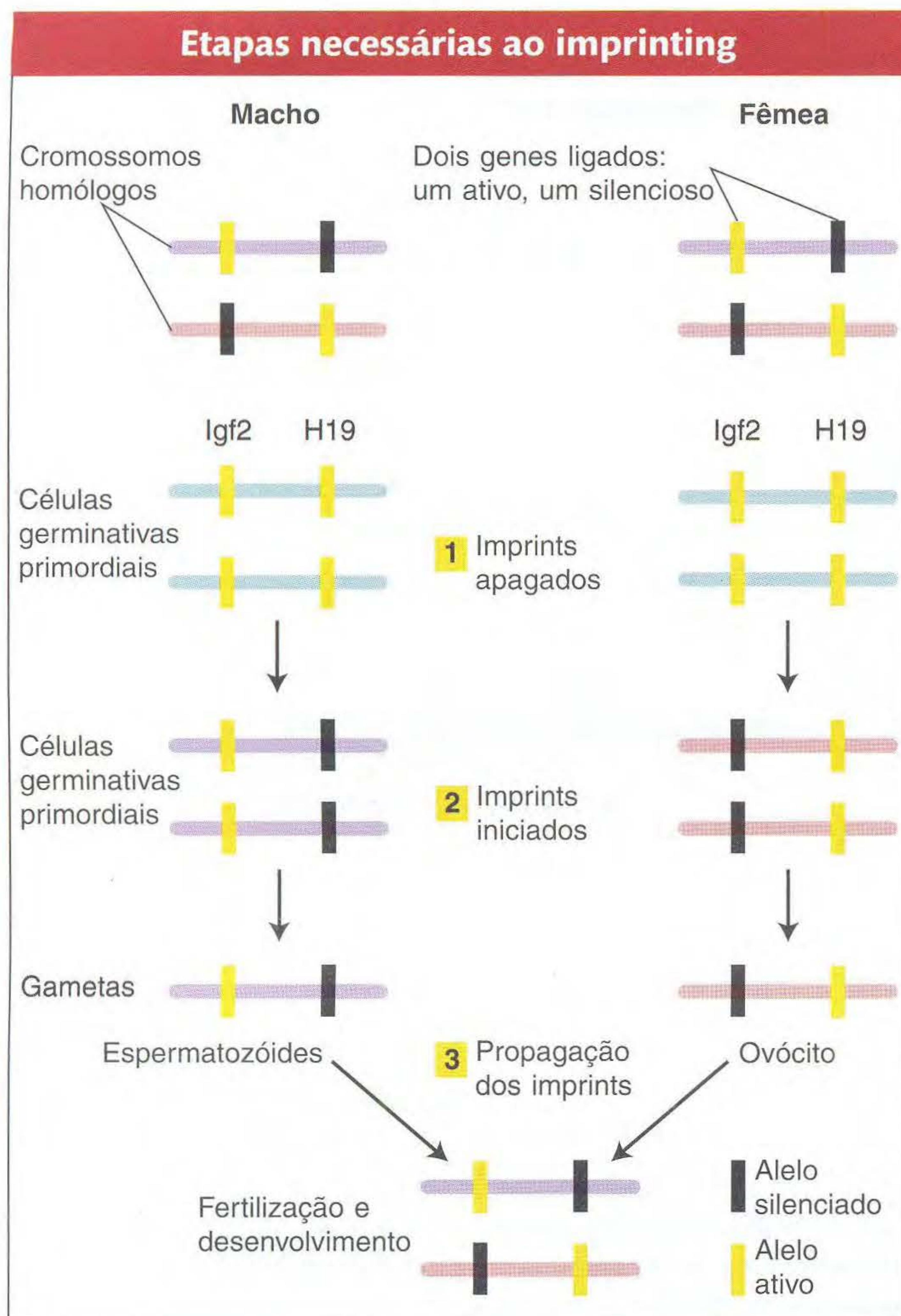


FIG. 11.21 Como *Igf2* e *H19* são imprintados diferencialmente em machos e fêmeas.

dos na maior parte do tempo. Vejamos agora essas grandes regiões do genoma que são transcricionalmente inativas. Um dos modelos mais úteis para compreender os mecanismos que mantêm a inatividade dos genes está relacionado ao controle do tipo reprodutivo em levedura e à mudança do tipo reprodutivo.

Mudança de tipo reprodutivo e silenciamento gênico

As células haplóides de leveduras são capazes de mudar seu tipo reprodutivo. As análises genética de alguns mutantes que não podiam mudar ou não podiam reproduzir-se (eles eram estéreis) foram fontes de importantes *insights* da mudança de tipo reprodutivo. Entre os mutantes de mudança estavam vários loci mutantes, incluindo o gene *HO* e os genes *HMRa* e *HMLα*. Estudos posteriores revelaram que o gene *HO* codifica uma endonuclease, uma enzima que corta o DNA, necessária para a iniciação da mudança. Também foi encontrado que os loci *HMRa* e *HMLα* contêm “cassetes” de informação genética não-expressa para os tipos reprodutivos *MATa* e *MATα*, respectivamente. Os loci *HMR* e *HML* são, portanto, chamados de cassetes “silenciosos”.

A endonuclease *HO* inicia a mudança de tipo reprodutivo inserindo uma quebra bifilamentar no locus *MAT*. A interconversão do tipo reprodutivo ocorre, então, por um tipo de recombinação entre o segmento de DNA (um cassette) de um dos dois loci não-expressos e o locus *MAT*. O resultado é a substituição do cassette antigo no locus *MAT* por um novo cassette. O tipo reprodutivo resultante é ou *MATa* ou *MATα*, dependendo de qual cassette está no locus *MAT* (Figura 11.22). O cassette inserido é, na verdade, copiado de *HML* ou do locus *HMR*. Desse modo, a mudança é reversível porque a informação para os cassetes *a* e *α* está sempre presente nos loci *HMR* e *HML* e nunca perdido.

Normalmente, os cassetes *HMR* e *HML* são “silenciosos”. Entretanto, nos mutantes *SIR* (reguladores de informação silenciosa), o silenciamento é comprometido de tal modo que tanto a informação de *a* quanto a de *α* são expressas. Os mutantes resultantes são estéreis. As proteínas *Sir2*, *Sir3* e *Sir4* formam um complexo que tem um papel-chave no silenciamento gênico. *Sir2* é uma histona desacetilase que facilita a condensação da cromatina e ajuda a ligar *HMR* e *HML* em domínios da cromatina que são inacessíveis a ativadores transcricionais. O **silenciamento gênico** é um processo muito diferente da repressão gênica; o silenciamento é um efeito de posição que depende da vizinhança na qual a informação genética está situada. Você aprenderá mais sobre os efeitos de posição mais adiante, na seção de variação do efeito de posição na mosca-das-frutas *Drosophila melanogaster*.

Em resumo, existem dois níveis distintos no controle do tipo reprodutivo de levedura. Primeiro, a regulação de um rearranjo do DNA controla a gama de produtos reguladores sintetizados dentro da célula. Segundo, as atividades de ligação ao DNA dessas proteínas reguladoras (*a1*, *α1* e *α2*) controlam as baterias de genes estruturais expressas dentro de cada tipo de célula. Esses dois níveis formam uma hierarquia: os genes do primeiro nível controlam a ativação dos genes do segundo nível, que, por sua vez, controlam a ativação dos genes estruturais. Esses genes estruturais codificam as proteínas que têm papéis no processo reprodutivo e a biologia de cada tipo celular. Como veremos com relação a animais no Capítulo 12, o controle genético do processo desenvolvimental em geral é hierárquico: redes de genes reguladores controlam a célula e a expressão histoespecífica das proteínas que medeiam o comportamento e função celulares.

Comparação da heterocromatina e eucromatina

Examinemos por que o silenciamento gênico, do tipo que silencia os cassetes *HML* e *HMR*, é um processo muito diferente da repressão gênica e qual o significado da vizinhança genômica. Para fazer isso, é importante notar que a cromatina não é uniforme em todos os cromossomos. Alguns domínios dos cromossomos têm uma cromatina altamente condensada, chamada de **heterocromatina**. Outros domínios são embalados em uma cromatina menos condensada, chamada de **eucromatina** (veja Figura 11.10b). A condensação da cromatina também muda no curso do ciclo celular. A cromatina das células que entram na mitose torna-se altamente condensada à medida que os cromossomos se alinham na preparação

A mudança de tipo reprodutivo é controlada pela recombinação de cassetes de DNA

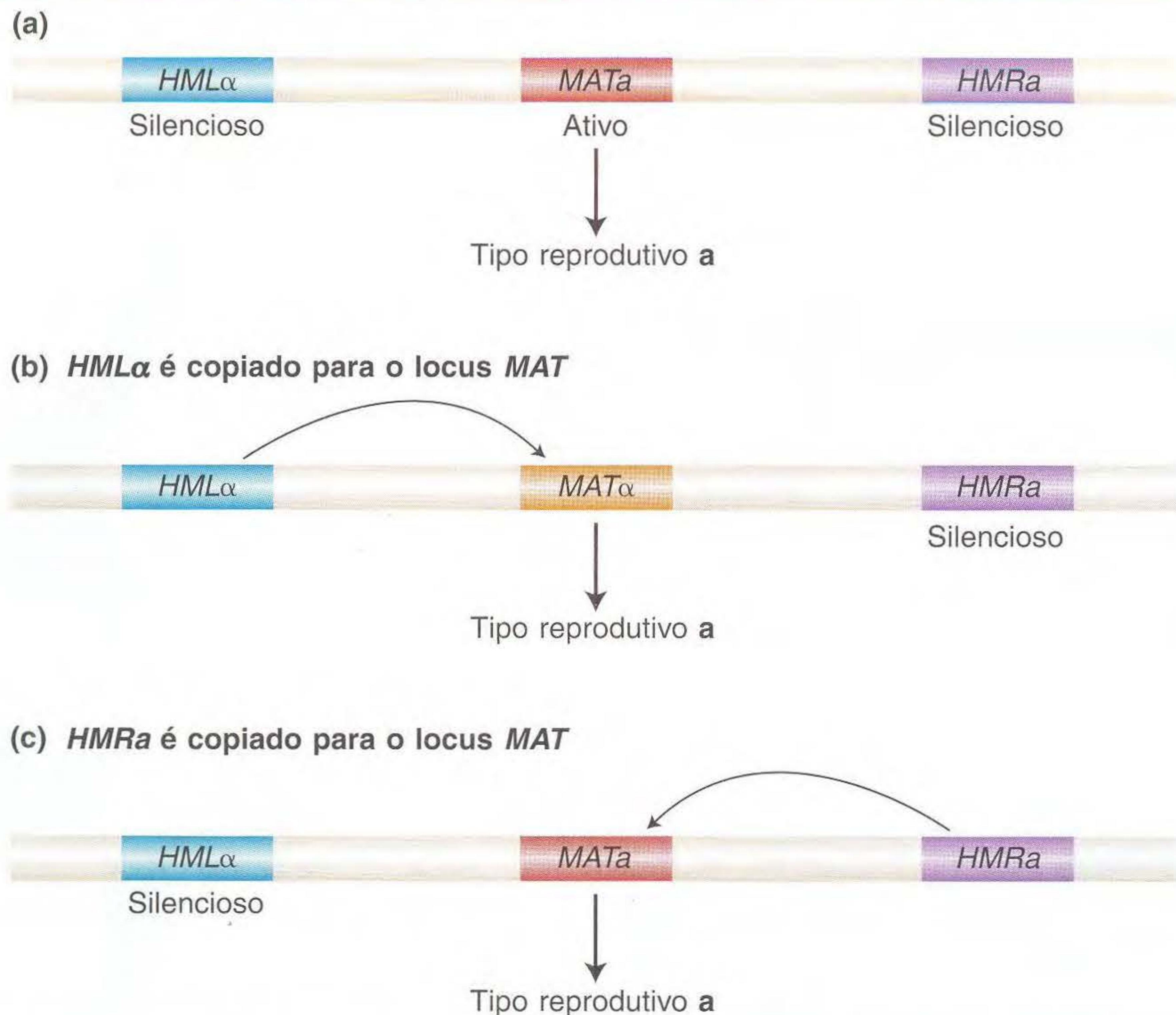


FIG. 11.22 O cromossomo III de *S. cerevisiae* codifica três loci de tipos reprodutivos, mas apenas os genes no locus MAT são expressos. HML codifica um cassette silencioso de genes α , e HMR codifica um cassette silencioso de genes a . A cópia de um cassette silencioso e a inserção por recombinação do locus MAT mudam o tipo reprodutivo.

para a divisão celular. Após a divisão celular, as regiões que formam a heterocromatina permanecem condensadas especialmente ao redor dos centrômeros e telômeros (chamada de **heterocromatina constitutiva**), enquanto as regiões que formam eucromatina tornam-se menos condensadas.

Os geneticistas primeiro suspeitaram de um papel limitado para a influência da estrutura da cromatina na regulação gênica no início da história da Genética. Naquela época, eles notaram que o DNA da heterocromatina continha poucos genes, enquanto a eucromatina era rica em genes. Mas o que é a heterocromatina se não genes? A maior parte do genoma eucariótico é composta de seqüências repetitivas que não fazem proteínas ou RNA estrutural, às vezes chamadas de DNA sucata (veja o Capítulo 14). Assim, os cromossomos densamente compactados de heterocromatina foram ditos formando uma estrutura “fechada” que era inacessível a proteínas reguladoras e não apropriada para a atividade gênica. Em contraste, a eucromatina, com seus nucleossomos mais amplamente espaçados, foi proposta como tendo uma estrutura “aberta” que permite a transcrição. A existência de regiões abertas e fechadas de cromatina também foi sugerida como um motivo para que as freqüências de recombinação sejam de 100 a 1.000 vezes maiores na eucromatina em comparação com a heterocromatina. A eucromatina, com sua conformação mais aberta, foi suposta como sendo mais acessível às proteínas necessárias para a recombinação do DNA.

A cromatina dos eucariontes não é uniforme. As regiões heterocromáticas altamente condensadas têm menos genes e menores freqüências de recombinação do que as regiões eucromáticas menos condensadas.

Variação do efeito de posição em *Drosophila* revela a vizinhança genômica

O geneticista Hermann Muller primeiro descobriu um fenômeno genético interessante enquanto estudava *Drosophila*: existem vizinhanças cromossômicas que podem silenciar genes que são experimentalmente “deslocados” para regiões adjacentes do cromossomo. Nesses experimentos, as moscas foram irradiadas com raios X para induzir mutações em suas células germinativas. A prole das moscas irradiadas foi triada quanto a fenótipos incomuns. Uma mutação no gene *white*, perto da ponta do cromossomo X, resultará na prole com olhos brancos em vez da cor vermelha tipo selvagem. Parte da prole tinha olhos muito incomuns, com manchas de cor branca e vermelha. Um exame citológico revelou um rearranjo cromossômico nas moscas mutantes: presente no cromossomo X estava uma inversão de um pedaço do cromossomo levando o gene *white* (Figura 11.23). As inversões e outros rearranjos cromossômicos serão discutidos no Capítulo 16. Nesse rearranjo, o gene *white*, que normalmente está situado em uma região eucromática do cromossomo X, agora se

O silenciamento gênico é causado pela expansão da heterocromatina

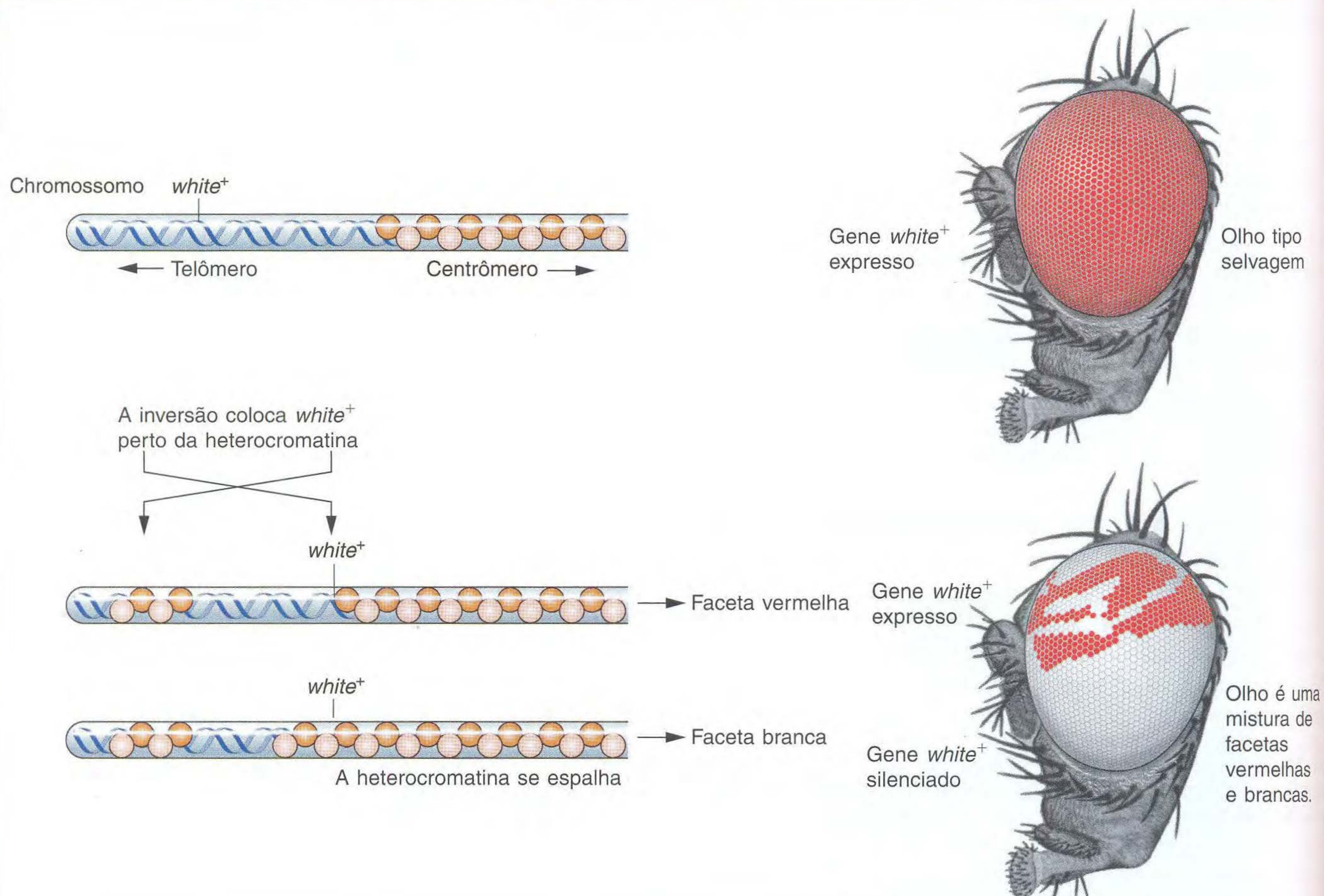


FIG. 11.23 O rearranjo cromossômico produz variação de efeito de posição. A inversão cromossômica coloca o alelo *white* tipo selvagem perto da heterocromatina. A dispersão da heterocromatina silencia o alelo. As facetas do olho são brancas em vez do vermelho tipo selvagem onde o alelo foi silenciado. [De C. Eissenberg and S. Elgin, *Encyclopedia of Life Sciences*. Nature Publishing Group, 2001, p. 3, Fig. 1.]

encontra perto do centrômero heterocromático. Em algumas células, a heterocromatina pode se “espalhar” para a eucromatina vizinha e silenciar o gene *white*. Manchas de tecido branco no olho são derivadas de descendentes de uma única célula na qual o gene *white* foi **epigeneticamente silenciado** e permanece silenciado ao longo de futuras divisões celulares. Em contraste, as manchas vermelhas surgem de células nas quais a heterocromatina não se espalhou para o gene *white*, e, assim, esse gene permanece ativo em todos os seus descendentes. A existência de manchas vermelhas e brancas nas células do olho de um único organismo ilustra marcadamente duas características do silenciamento epigenético. Primeiro, a expressão de um gene pode ser reprimida em virtude de sua posição no cromossomo em vez de uma mutação em sua sequência de DNA. Segundo, o silenciamento epigenético pode ser herdado de uma geração celular para a seguinte.

Os achados de estudos subsequentes em *Drosophila* e leveduras demonstraram que muitos genes ativos são silenciados de modo em mosaico quando deslocados para a vizinhança (perto de centrômeros ou telômeros) que é constituída de heterocromáticos. Assim, a habilidade da heterocromatina de se espalhar para a eucromatina e silenciar genes é uma característica comum a muitos organismos. Esse fenômeno foi chamado de **variação de efeito de posição (PEV)**. Ele dá uma poderosa evidência de que a estrutura da cromatina é capaz de regular a expressão dos genes, nesse caso, determinando se os genes com sequência idêntica de DNA serão ativos ou silenciados.

Os genes ativos deslocados para vizinhanças genômicas que são heterocromáticas podem ser silenciados se a heterocromatina se espalhar para os genes.

A análise genética de PEV revela proteínas necessárias para a formação de heterocromatina

Para descobrir que proteínas podem estar implicadas no estabelecimento da heterocromatina, os geneticistas isolaram mutações em um segundo locus cromossômico que suprimia ou acentuava o padrão variegado (Figura 11.24). Os supressores da variação [chamados de *Su(var)*] são genes

Alguns genes acentuam ou suprimem o espalhamento da heterocromatina

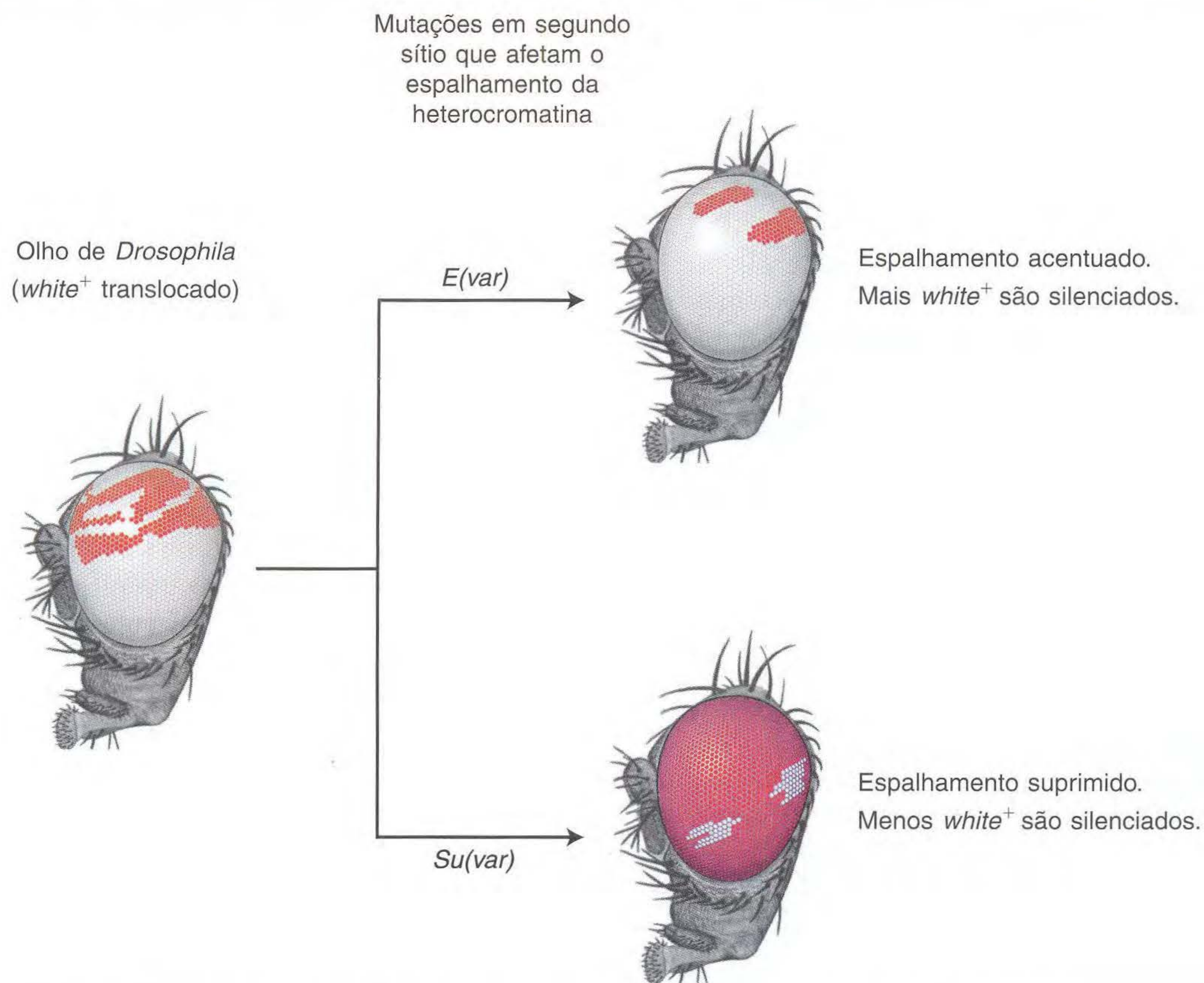
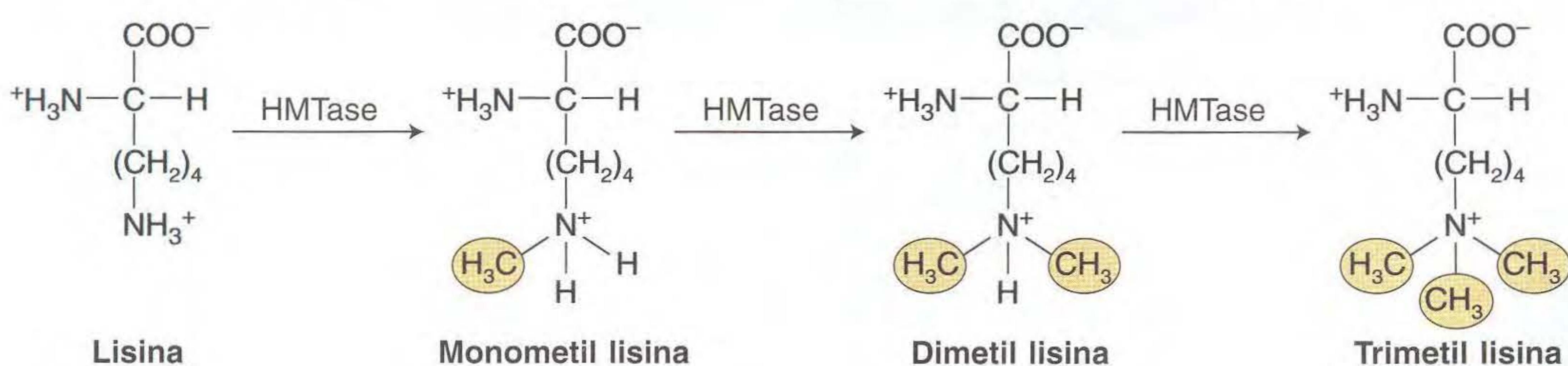


FIG. 11.24 Mutações foram usadas para identificar genes que suprimem, *Su(var)*, ou acentuam, *E(var)*, variação de efeito de posição. [De J. C. Essenberg and S. Elgin, *Encyclopedia of Life Sciences*. Nature Publishing Group, 2001, p. 3, Fig. 1.]

que, quando mutados, reduzem a dispersão da heterocromatina, significando que os produtos tipo selvagem desses genes são necessários para a dispersão. De fato, os alelos *Su(var)* provaram ser um tesouro para os cientistas interessados nas proteínas que são necessárias para estabelecer e manter o estado inativo da heterocromatina. Entre mais de 50 produtos gênicos de *Drosophila* identificados por essa triagem, estava a **proteína-1 de heterocromatina (HP-1)** que previamente tinha sido associada a telômeros e centrômeros heterocromáticos. Assim, faz sentido que uma mutação no gene que codifica HP-1 apresente-se como um alelo *Su(var)* porque a proteína é necessária, de algum modo, para produzir ou manter a heterocromatina. Outro gene *Su(var)* foi encontrado codificando uma metiltransferase que adiciona grupos metil a um aminoácido específico na cauda de histona H3 (chamada de metiltransferase H3 ou HMTase). Uma das reações catalisadas pela HMTase é mostrada aqui:

Proteínas similares a HP-1 e HMTase foram isoladas em diversos seres, sugerindo a conservação de uma importante função eucariótica.

Vimos que as regiões ativamente transcritas estão associadas a nucleossomos cujas caudas de histona são hiperacetiladas, e que ativadores transcripcionais, tais como *GCN5*, codificam uma atividade de histona acetiltransferase. Como discutido daqui em diante, as marcas de acetil também podem ser removidas das histonas pelas histonas desacetilase. Similarmente, a cromatina feita de nucleossomos que são metilados na lisina 9 de H3 (chamado H3meK9) são ligadas à proteína HP-1 que contém marcas epigenéticas que estão associadas à heterocromatina. Os cientistas agora são capazes de separar heterocromatina e eucromatina e analisar as diferenças nas modificações de histonas e proteínas ligadas. O procedimento usado, a imunoprecipitação de cromatina (ChIP), é descrito no Capítulo 20.



A heterocromatina pode se espalhar mais em algumas células que em outras

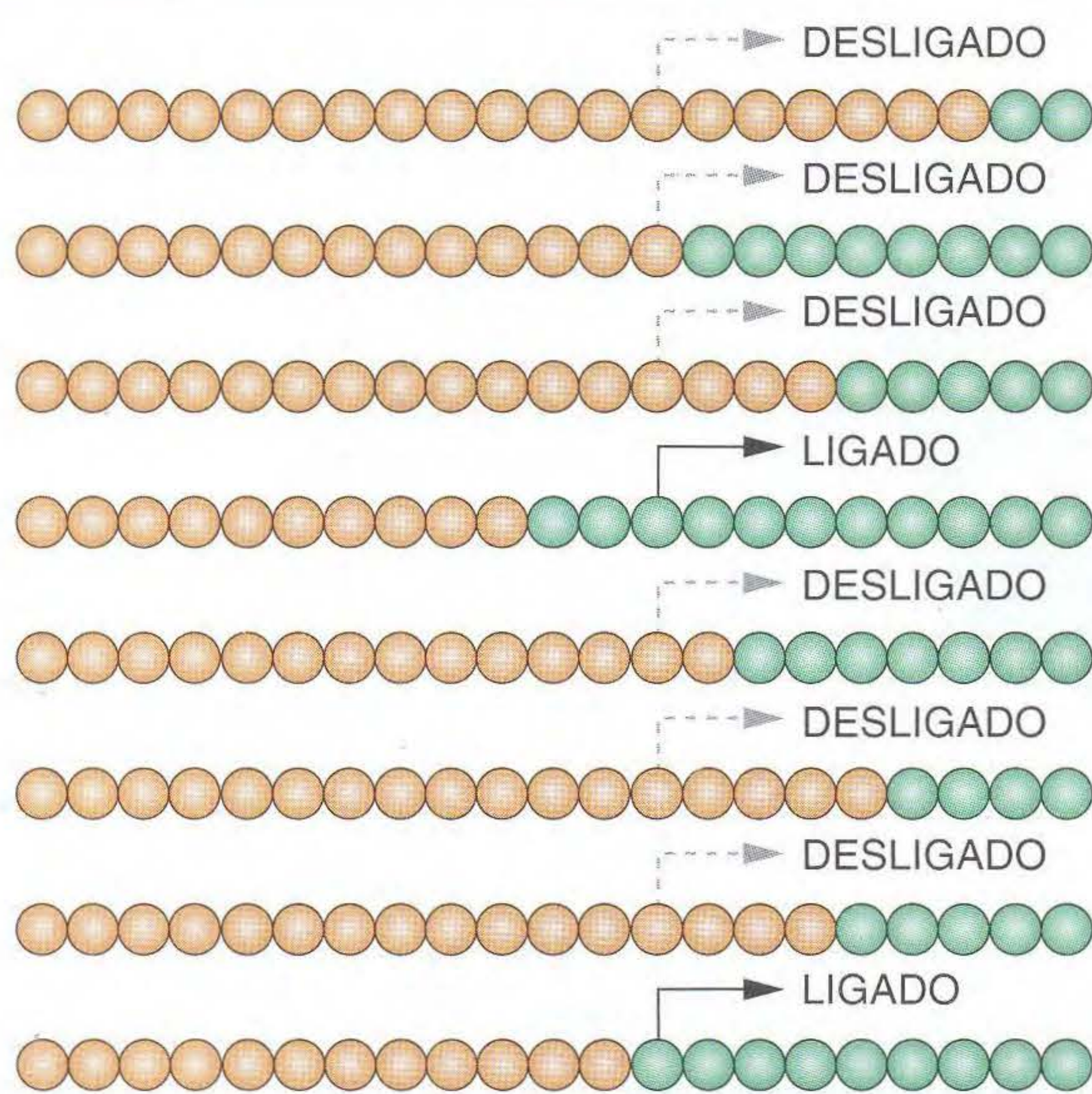


FIG. 11.25 A dispersão da heterocromatina na eucromatina adjacente é variável. Em quatro células diploides geneticamente idênticas, a heterocromatina espalha-se o suficiente para desligar um gene em alguns cromossomos, mas não em outros. A heterocromatina e a eucromatina são representadas por esferas laranja e verde, respectivamente. [De M. Gaszner and G. Felsenfeld, "Insulators: Exploiting Transcriptional and Epigenetic Mechanisms," *Nat. Rev. Genet.* 7, 2006, 703-713.]

A Figura 11.25 ilustra como, na ausência de qualquer barreira, a heterocromatina pode se espalhar para as regiões vizinhas em algumas células mas não em outras e inativar genes. Poderia ser o que acontece com o gene *white* de *Drosophila* quando ele é translocado para perto do domínio de heterocromatina associado às pontas dos cromossomos. Mas o espalhamento da heterocromatina pode ser parado? Podemos imaginar que a dispersão da heterocromatina para regiões com atividade gênica pode ser desastrosa para um organismo porque os genes ativos seriam silenciados à medida que são convertidos em heterocromatina. Para evitar esse desastre potencial, um genoma contém elementos de DNA chamados de **barreiras**

de isolamento que impedem a dispersão da heterocromatina criando um ambiente local que não é favorável à formação de heterocromatina. Por exemplo, uma barreira de isolamento poderia ligar HAT e, assim fazendo, com certeza as histonas adjacentes seriam hiperacetiladas. Um modelo de como uma barreira de isolamento pode agir para "proteger" uma região de eucromatina de ser convertida em heterocromatina é mostrado na Figura 11.26.

Silenciamento de um cromossomo inteiro: inativação do cromossomo X

Um fenômeno epigenético chamado de inativação do cromossomo X tem intrigado os cientistas por décadas. No Capítulo 16, você aprenderá sobre os efeitos do número de cópias gênicas no fenótipo de um organismo. Por enquanto é suficiente saber que o número de transcritos produzidos por um gene geralmente é proporcional ao número de cópias desse gene em uma célula. Os mamíferos, por exemplo, são diplóides e têm duas cópias de cada gene situadas em seus autossomos. Para a grande maioria dos genes, ambos os alelos são expressos. Entretanto, isso não é possível para os cromossomos sexuais. Como discutido no Capítulo 2, o número de cromossomos X e Y difere entre os sexos, com as fêmeas de mamíferos tendo dois cromossomos X e os machos tendo apenas um. O cromossomo X de mamíferos é tido como contendo cerca de 1.000 genes. As fêmeas têm o dobro de cópias dos genes ligados ao X, e, de outro modo, expressariam duas vezes mais transcritos desses genes que os machos se não houvesse um mecanismo para corrigir esse desequilíbrio. (Não ter um cromossomo Y não é um problema para as fêmeas, porque muito poucos genes nesse cromossomo são necessários apenas para o desenvolvimento dos machos.) Esse desequilíbrio de dosagem é corrigido por um processo chamado de **compensação de dose**, que torna a quantidade da maioria dos produtos gênicos das duas cópias do cromossomo X nas fêmeas equivalente à dose única de cromossomo X nos machos. Em mamíferos, essa equivalência é obtida pela inativação aleatória de um dos dois cromossomos X em cada célula em um estágio inicial no desenvolvimento. Esse estado inativo é então propagado para todas as células da prole. (Na linhagem

Os isoladores de barreiras param o espalhamento da heterocromatina

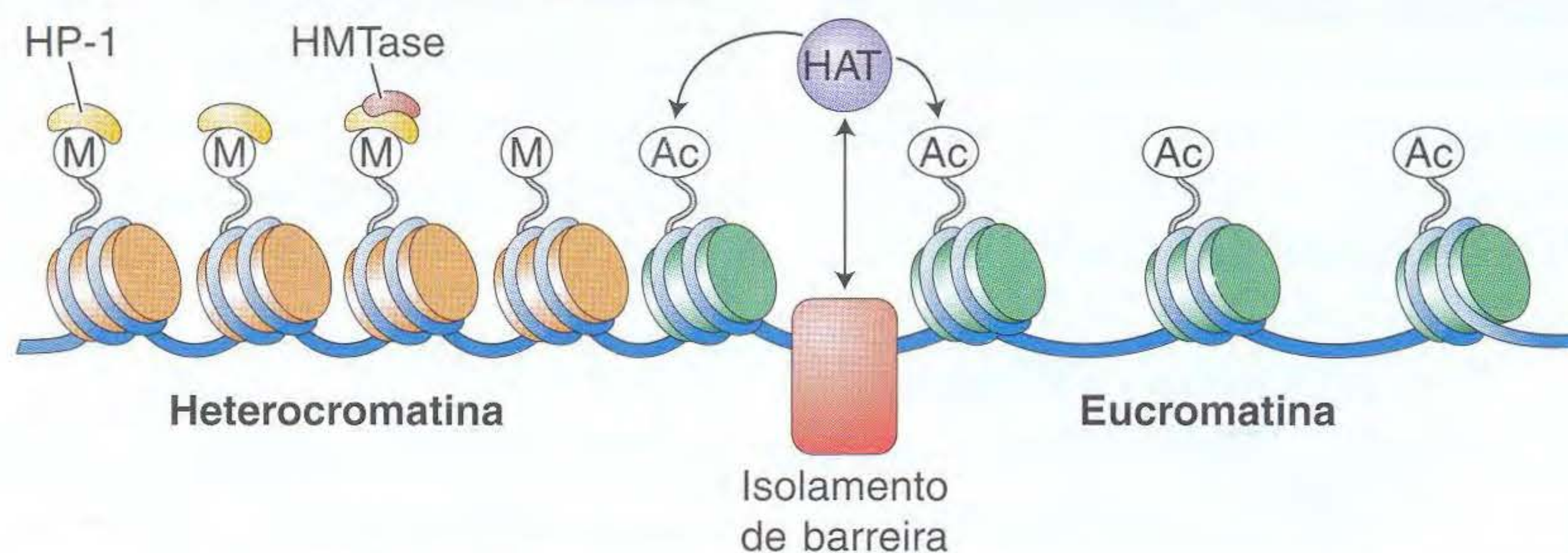


FIG. 11.26 Nesse modelo, os isoladores de barreiras recrutam atividades enzimáticas, tais como histona acetiltransferase (HAT), que promovem a formação de eucromatina. A letra "M" representa metilação, e as letras "Ac", acetilação. [De Gaszner and G. Felsenfeld, "Insulators: Exploiting Transcriptional and Epigenetic Mechanisms," *Nat. Rev. Genet.* 7, 2006, 703-713.]

Mecanismos diferentes de compensação de dose

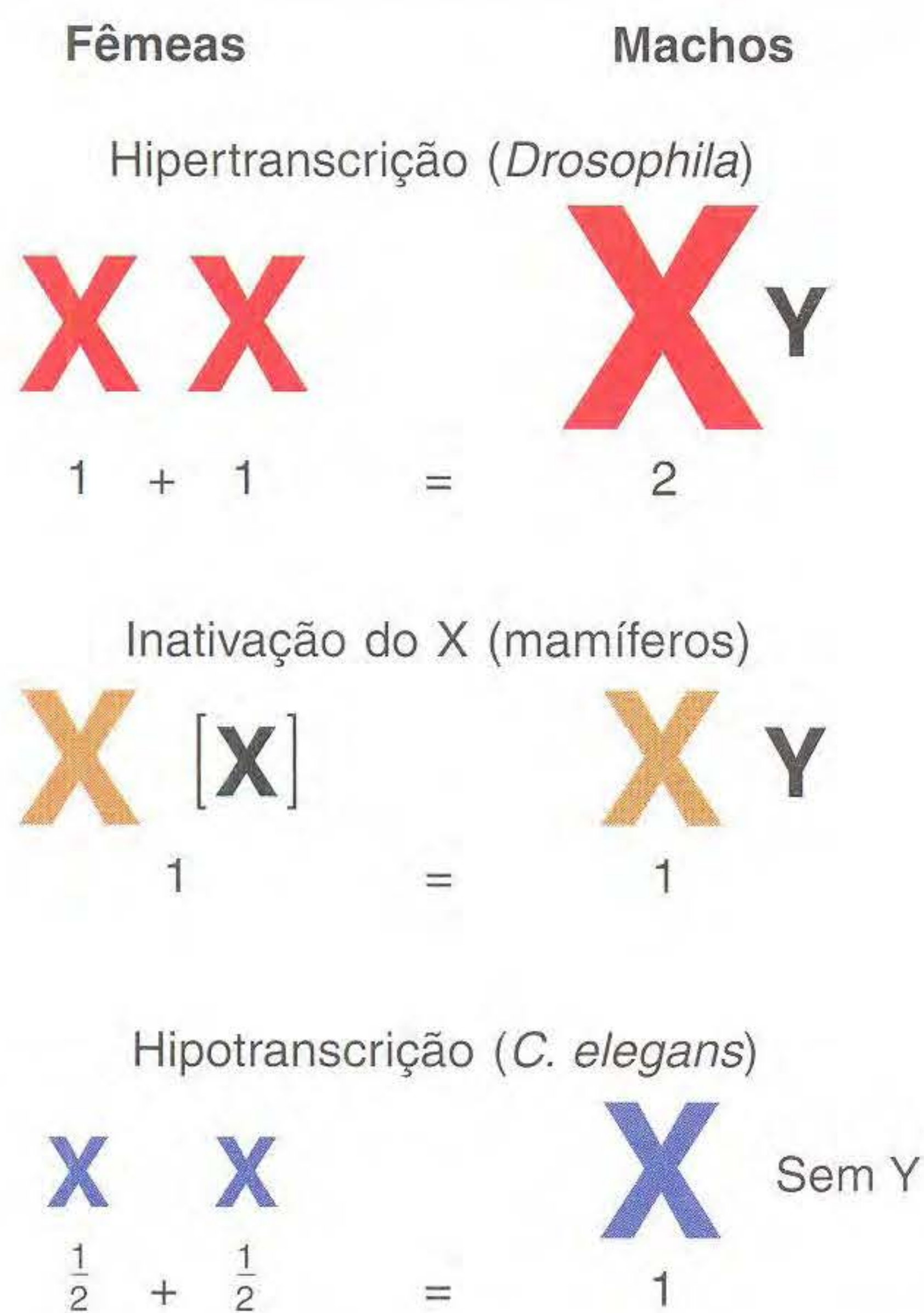


FIG. 11.27 A compensação de dose pode ser obtida duplicando a expressão do cromossomo X masculino (hipertranscrição), pela inativação do X, ou reduzindo à metade a expressão de ambos os cromossomos X das fêmeas (hipotranscrição).

germinativa, o segundo cromossomo X torna-se reativado na ovocitogênese.) O cromossomo inativado, chamado de **corpúsculo de Barr**, pode ser visto no núcleo como uma estrutura heterocromática fortemente corada, altamente condensada.

Dois aspectos da inativação do cromossomo X são relevantes para uma discussão da cromatina e a regulação da expressão gênica. Primeiro, a maioria dos genes no cromossomo X inativado são silenciados, e o cromossomo tem marcas epigenéticas associadas à heterocromatina, incluindo a metilação de H3 na lisina 9 e hipermetilação de seu DNA. Segundo, os genes no cromossomo inativo permanecem inativos em todas as descendentes dessas células. Como a própria sequência do DNA não é alterada, essa alteração herdável é um exemplo de herança epigenética.

Curiosamente, embora vários grupos exibam compensação de dose, o mecanismo de compensação pode diferir acentuadamente. Por exemplo, nas moscas-das-frutas, a expressão de genes no cromossomo X é compensada não pela inativação de um dos X das fêmeas, mas duplicando a expressão de genes no cromossomo X do macho (Figura 11.27). Esse mecanismo é caracterizado pela ligação de um complexo RNA-proteína, chamada MSL, juntamente com todo o tamanho do cromossomo X em homens (veja ilustração na abertura do capítulo). Um dos componentes do complexo MSL é uma histona acetiltransferase. Lembre que as histonas acetiladas são uma característica importante da cromatina ativa. Assim, a função do complexo

MSL parece ser adicionar grupos acetil às histonas. MSL vem de *male-specific lethal* (letal específico de machos), e o complexo foi assim chamado porque as triagens genéticas para mutações letais em machos identificaram seus componentes.

Para a maioria dos organismos diplóides, ambos os alelos de um gene são expressos independentemente. A inativação do X e o imprinting genômico são exemplos de expressão monoalélica. Nesses casos, os mecanismos epigenéticos silenciam uma cópia de um cromossomo inteiro ou de um único locus cromossômico, respectivamente.

A herança de marcas epigenéticas e estrutura cromossômica

A herança epigenética pode ser definida operacionalmente como a herança de estados da cromatina de uma geração celular para a seguinte. O que essa herança significa é que, na replicação do DNA, tanto a sequência do DNA quanto a estrutura da cromatina são fielmente passadas para a geração seguinte. Entretanto, ao contrário da sequência do DNA, a estrutura da cromatina pode mudar no curso do ciclo celular quando, por exemplo, os fatores de transcrição modificam o código de histonas, causando mudanças locais na posição do nucleossomo e/ou densidade do nucleossomo.

Como mencionado no Capítulo 7, o replissomo não só copia os filamentos parentais, mas também desmonta os nucleossomos nos filamentos parentais e os reorganiza tanto nos filamentos parentais quanto nos filhos. Esse processo é feito pela distribuição aleatória das histonas antigas de nucleossomos existentes para os filamentos-filhos e pela entrega de novas histonas ao replissomo. Desse modo, as histonas antigas com suas caudas modificadas e as novas histonas com caudas não-modificadas são montadas em nucleossomos que se tornam associados a ambos os filamentos-filhos. O código levado pelas histonas antigas mais provavelmente orienta a modificação das novas histonas (Figura 11.28).

A herança da metilação do DNA é mais bem compreendida. A replicação semiconservativa gera hélices-filhas que são metiladas em um de seus dois filamentos (o filamento parental). Os filamentos não-metilados são metilados pelas DNA metiltransferases que têm alta afinidade por esses chamados de substratos **hemimetilados** e são orientados pelo padrão de metilação do filamento parental (Figura 11.29). Assim, a informação inerente ao código de histona e os padrões existentes de metilação do DNA servem para reconstituir a estrutura local de cromatina que existia antes da síntese de DNA e mitose.

A estrutura da cromatina é herdada de uma geração celular para outra porque existem mecanismos para replicar o DNA juntamente com marcas epigenéticas associadas.

Herança de estados da cromatina

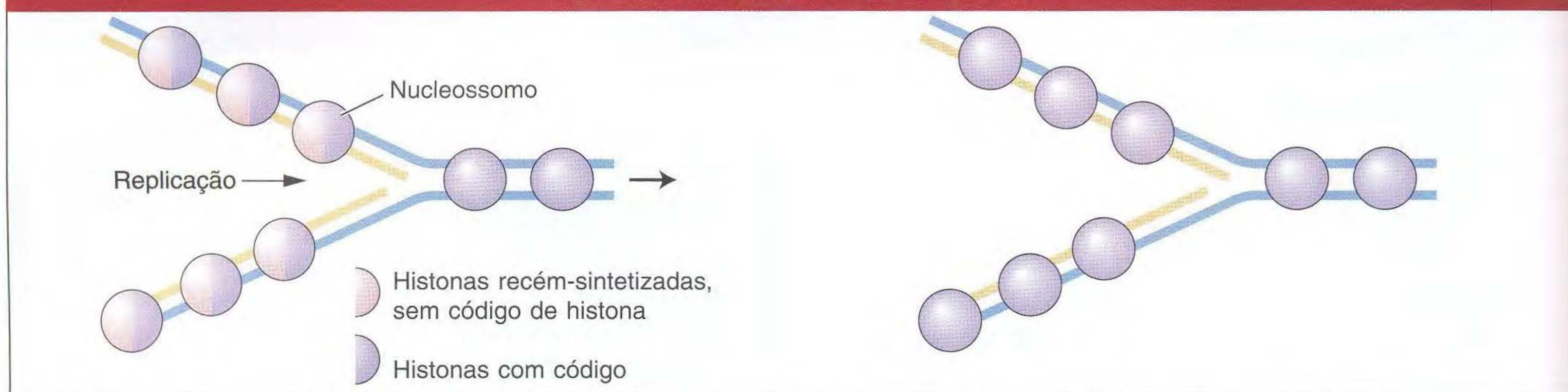


FIG. 11.28 Na replicação, as histonas antigas (púrpura) com seus códigos de histona são distribuídas aleatoriamente para os filamentos-filhos, onde dirigem o código de histonas adjacentes recém-montadas (laranja) para formar nucleossomos completos.

Um modelo de herança da metilação do DNA

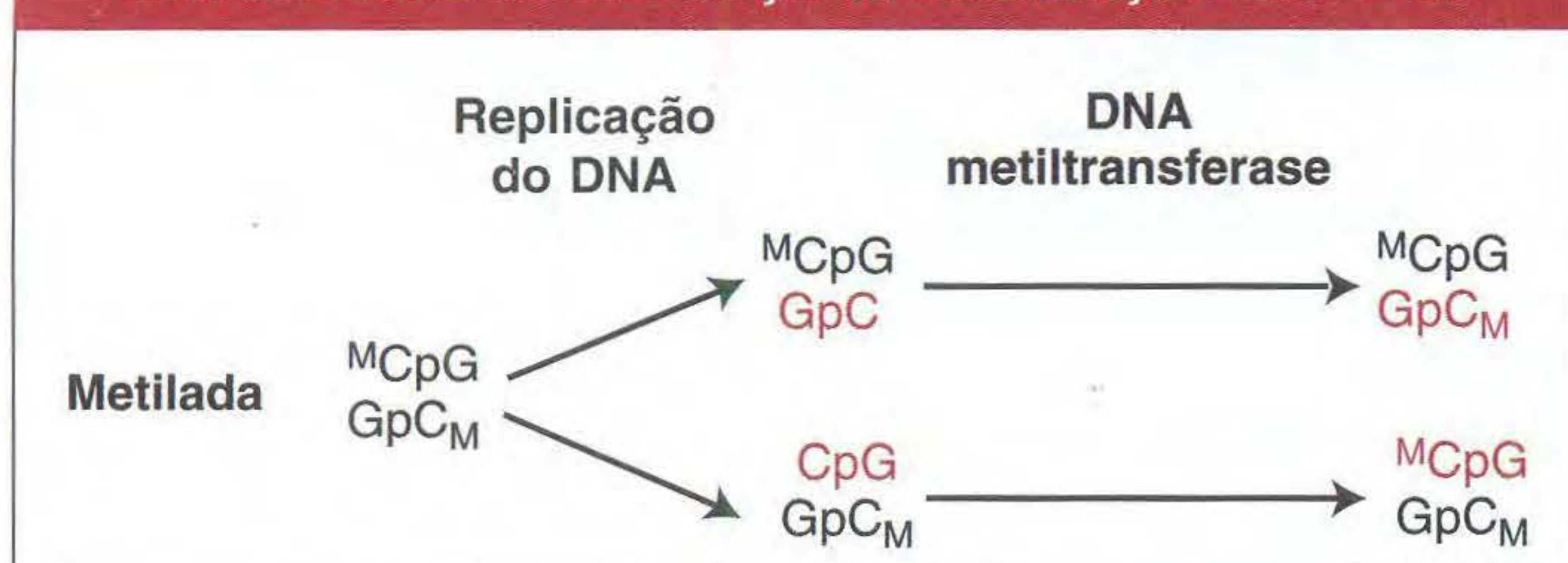


FIG. 11.29 Após a replicação, os dinucleotídeos CG hemimetilados (mostrados como CpG) são totalmente metilados. Os filamentos parentais são pretos, e os filamentos-filhos são vermelhos. A letra "M" representa o grupo metil no nucleotídeo C. [De Y. H. Jiang, J. Bressler, and A. L. Beaudet, "Epigenetics and Human Disease," *Annu. Rev. Genomics Hum. Genet.* 5, 2004, 479-510.]

► Resumo

Muitos aspectos da regulação gênica eucariótica se assemelham à regulação de óperons bacterianos. Ambos operam amplamente ao nível da transcrição, e ambos são baseados em proteínas de ação trans que se ligam a seqüências-alvo reguladoras de ação cis na molécula de DNA. Essas proteínas regulatórias determinam o nível de transcrição de um gene controlando a ligação da RNA polimerase ao promotor do gene.

Existem três categorias distintivas principais do controle da transcrição em eucariontes. Primeira, os genes eucarióticos possuem acentuadores, que são elementos reguladores de ação cis situados, às vezes, a grandes distâncias lineares do promotor. Muitos genes possuem vários acentuadores. Segunda, esses acentuadores são, em geral, ligados por mais fatores de transcrição que os óperons bacterianos. Os eucariontes multicelulares podem gerar milhares de padrões da expressão gênica com um número limitado de proteínas reguladoras (fatores de transcrição). Eles fazem isso por interações combinatórias entre os fatores de transcrição. Os acentuassomos são complexos de proteínas reguladoras que interagem de modo cooperativo e sinérgico para promover altos níveis de transcrição pelo recrutamento da RNA polimerase II para o sítio de início da transcrição.

Terceira, os genes eucarióticos são embalados em cromatina. A ação gênica e a repressão requerem modificações específicas da cromatina. A grande maioria das dezenas de milhares de genes em um genoma tipicamente eucariótico estão desligados em qualquer momento. Os genes são mantidos em um estado transcricionalmente inativo pela partici-

pação de nucleossomos, que servem para compactar a cromatina e evitar a ligação da RNA polimerase II. A posição dos nucleossomos e a extensão da condensação de cromatina são instruídas pelo código de histonas, o padrão de modificações pós-traducionais das caudas de histona. O código de histona é uma marca epigenética que, juntamente com a metilação de bases citosina, pode ser alterado por fatores de transcrição. Esses fatores ligam-se a regiões reguladoras e recrutam complexos proteicos que modificam enzimaticamente nucleossomos adjacentes. Esses grandes complexos proteicos multissubunitários usam energia da hidrólise de ATP para mover nucleossomos e remodelar a cromatina.

A existência de fenômenos epigenéticos tais como o imprinting genético e a inativação de cromossomos X demonstra que a expressão gênica eucariótica pode ser silenciada sem mudar a seqüência de DNA do gene. Outro fenômeno epigenético, variação do efeito de posição, revelou a existência de domínios heterocromáticos repressivos que estão associados a nucleossomos altamente condensados e contêm poucos genes. As barreiras isoladoras mantêm a integridade do genoma, impedindo a conversão da eucromatina em heterocromatina.

A replicação do DNA copia fielmente tanto a seqüência do DNA quanto a estrutura da cromatina das células genitoras para as células-filhas. As células recém-formadas herdam tanto a informação genética, inerente à seqüência de nucleotídeos do DNA, quanto a informação epigenética, que está no código de histonas e o padrão de metilação do DNA.

Termos-chave

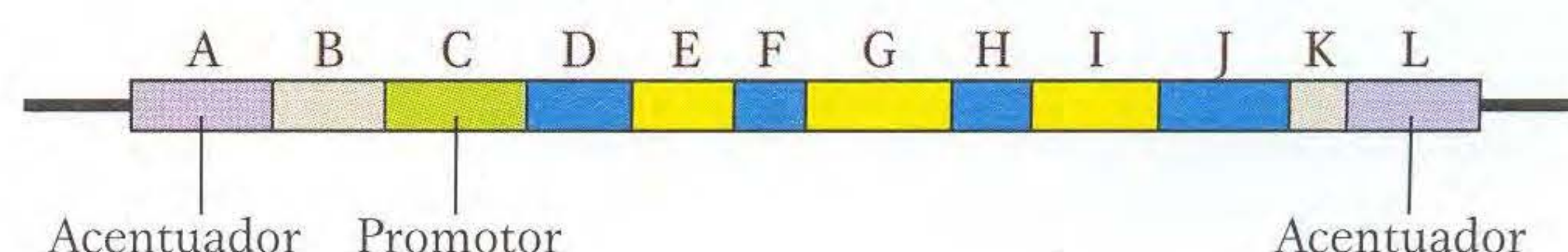
acentuador	eucromatina	imprinting paterno
acentuassomo	feromônio	isolador de bloqueio de acentuador
barreira de isolamento	gene repórter	marca epigenética
cauda de histona	hemimetilação	metilação do DNA
coativador	herança epigenética	proteína-1 de heterocromatina (HP-1)
código de histona	herança monoalélica	remodelagem de cromatina
compensação de dose	heterocromatina	seqüência de ativação antecedente (UAS)
complexo mediador	heterocromatina constitutiva	silenciamento epigenético
co-repressor	hiperacetilação	silenciamento gênico
corpúsculo de Barr	hipoacetilação	variação de efeito de posição (PEV)
domínio de ativação	histona desacetilase (HDAT)	
efeito sinérgico	imprinting genômico	
elemento proximal do promotor	imprinting materno	

Problemas

PROBLEMAS BÁSICOS

1. Que analogias você pode fazer entre os fatores transcripcionais de ação trans que ativam a expressão gênica em eucariontes e os fatores correspondentes em bactérias? Cite um exemplo.
2. Contraste os estados dos genes em bactérias e eucariontes com relação à ativação gênica.
3. Preveja e explique o efeito na transcrição de *GAL1*, na presença apenas de galactose, das seguintes mutações:
 - a. Deleção de um sítio de ligação de Gal4 no elemento UAS *GAL1*.
 - b. Deleção de todos os quatro sítios de ligação de Gal4 no elemento UAS *GAL1*.
 - c. Deleção do sítio de ligação Mig1 antecedente de *GAL1*.
 - d. Deleção do domínio de ativação Gal4.
 - e. Deleção do gene *GAL80*.
 - f. Deleção do promotor *GAL1*.
 - g. Deleção do gene *GAL3*.
4. Como a ativação do gene *GAL1* é ativada na presença de galactose e glicose?
5. Quais são os papéis da desacetilação e acetilação de histona na regulação gênica, respectivamente?
6. Foi isolada uma linhagem α de levedura que não pode mudar o tipo reprodutivo. Que mutações ela pode levar que expliquem esse fenótipo?
7. Que genes são regulados pelas proteínas $\alpha 1$ e $\alpha 2$ em uma célula α ?
8. O que são proteínas Sir? Como as mutações nos genes *SIR* afetam a expressão dos cassetes de tipo reprodutivo?
9. O que significa o termo *herança epigenética*? Cite dois exemplos de tal herança.
10. O que é um acentuassomo? Por que uma mutação em qualquer uma das proteínas acentuassomo reduziria muito a taxa de transcrição?
11. Por que as mutações nos genes imprintados geralmente são dominantes?
12. Que características distinguem um gene silenciado epigeneticamente de um gene que não é expresso, devido a uma alteração em sua seqüência de DNA?

13. Que mecanismos são tidos como responsáveis pela herança da informação epigenética?
14. Qual a diferença fundamental em como são regulados os genes bacterianos e eucarióticos?
15. Por que se diz que a regulação transricional em eucariontes é caracterizada por interações combinatórias?
16. O diagrama seguinte representa a estrutura de um gene em *Drosophila melanogaster*; os segmentos azuis são éxons, e os segmentos amarelos, íntrons.



- a. Que segmentos do gene serão representados no RNA transcrito inicial?
- b. Que segmentos do gene serão removidos pela recomposição do RNA?
- c. Que segmentos mais provavelmente irão ligar-se a proteínas que interagem com a RNA polimerase?

PROBLEMAS DESAFIADORES

17. A transcrição de um gene chamado *YFG* (your favorite gene) é ativada quando três fatores de transcrição (TFA, TFB, TFC) interagem para recrutar o coativador CRX. TFA, TFB, TFC e CRX e seus respectivos sítios de ligação constituem um acentuassomo situado a 10 kb do ponto de início da transcrição. Faça um diagrama mostrando como você acha que o acentuassomo funciona para recrutar a RNA polimerase para o promotor de *YFG*.
18. Uma única mutação em um dos fatores de transcrição no Problema 17 resulta em uma drástica redução na transcrição de *YFG*. Faça um diagrama para representar essa interação de mutantes.
19. Diagrame o efeito de uma mutação no sítio de ligação para um dos fatores de transcrição no Problema 17.
20. Como um gene epigeneticamente silenciado difere de um gene mutante (um alelo nulo do mesmo gene)?

21. O que são marcas epigenéticas? Quais estão associadas à heterocromatina? Como as marcas epigenéticas são interpretadas na estrutura da cromatina?
22. Você recebe quatro linhagens de leveduras pelo correio, e as instruções acompanhantes dizem que cada linhagem contém uma única cópia do transgene A. Você cultiva as quatro linhagens e determina que apenas três linhagens expressam o produto proteico do transgene A. Análises posteriores revelam que o transgene A está situado em uma posição no genoma de levedura em cada uma das quatro linhagens. Forneça uma hipótese para explicar esse resultado.
23. Em *Neurospora*, todos os mutantes que afetam as enzimas carbamil fosfato sintetase e aspartato transcarbamilase estão mapeados no locus *pyr-3*. Se você induz mutações *pyr-3* por ICR-170 (um mutágeno químico), você descobre que ou ambas as funções enzimáticas estão faltando, ou apenas a função transcarbamilase está faltando; em nenhum caso a atividade de sintetase está faltando quando a atividade de transcarbamilase está presente. (ICR-170 supostamente induz mudanças de matriz de leitura.) Interprete esses resultados com relação a um possível óperon.
24. Você quer descobrir os elementos reguladores de ação cis de DNA responsáveis pelas respostas transcricionais de dois genes, *c-fos* e *globina*. A transcrição do gene *c-fos* é ativada em resposta ao fator de crescimento de fibroblasto (FGF), mas é inibida pelo cortisol (Cort). Por outro lado, a transcrição do gene de *globina* não é afetada por FGF ou cortisol, mas é estimulada pelo hormônio eritropoietina (EP). Para descobrir os elementos de DNA de regulação cis responsáveis por essas respostas transcricionais, você usa os seguintes clones dos genes *c-fos* e *globina*, bem como duas combinações “híbridas” (genes de fusão), como mostrado no diagrama 1. A letra A representa o gene *c-fos* intacto, D representa o gene *globina* intacto, e B e C representam os genes de fusão *c-fos-globina*. Os éxons *c-fos* e *globina* (E) e os íntrons (I) são numerados. Por exemplo, E3(f) é o terceiro éxon do gene *c-fos* e I2(g) é o segundo íntron do gene de *globina*. (Essas marcas são dadas para ajudá-lo a dar uma resposta clara.) São indicados os sítios de início da transcrição (setas pretas) e os sítios de poliademilação (setas vermelhas).

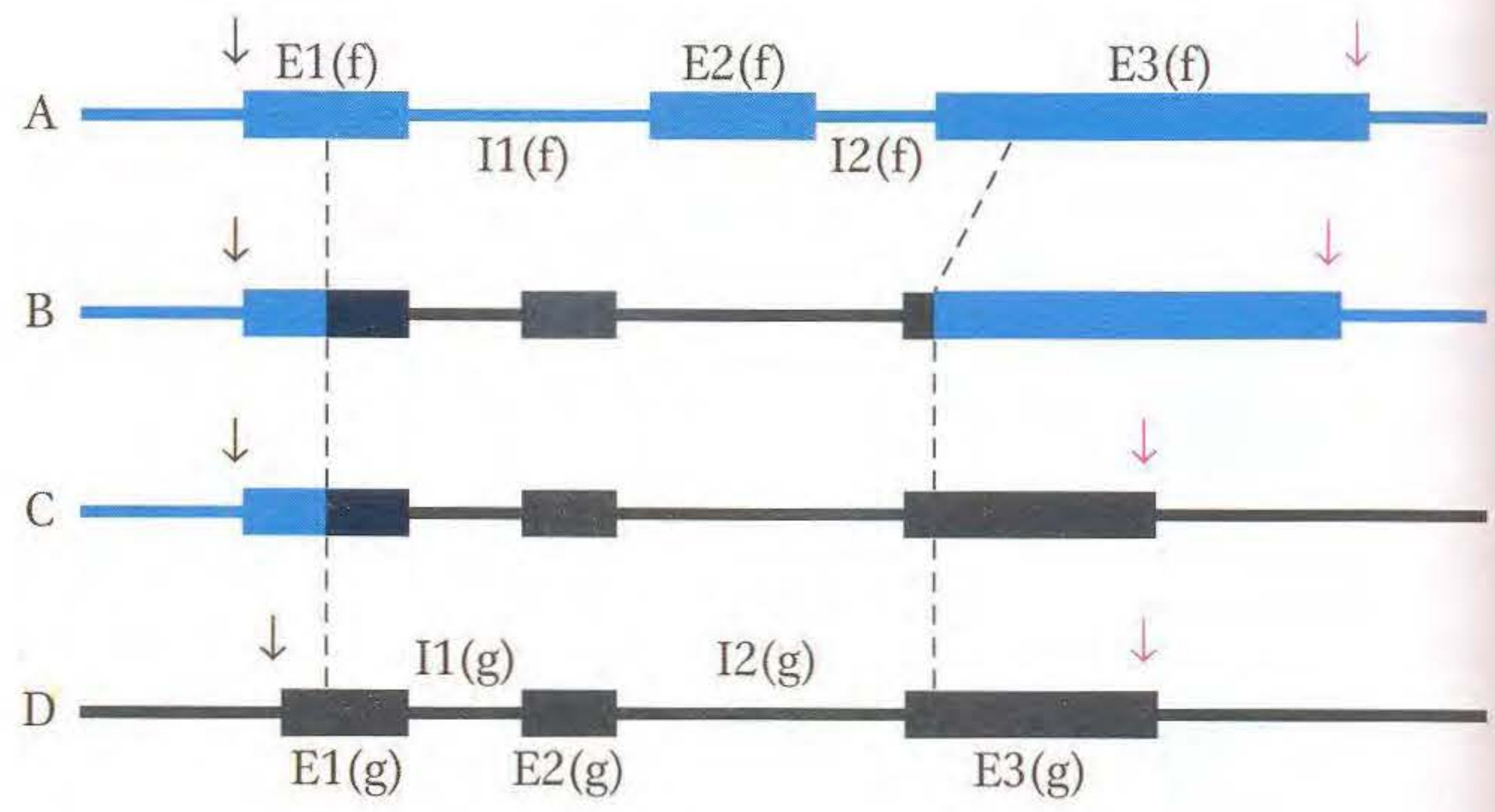


Diagrama 1.

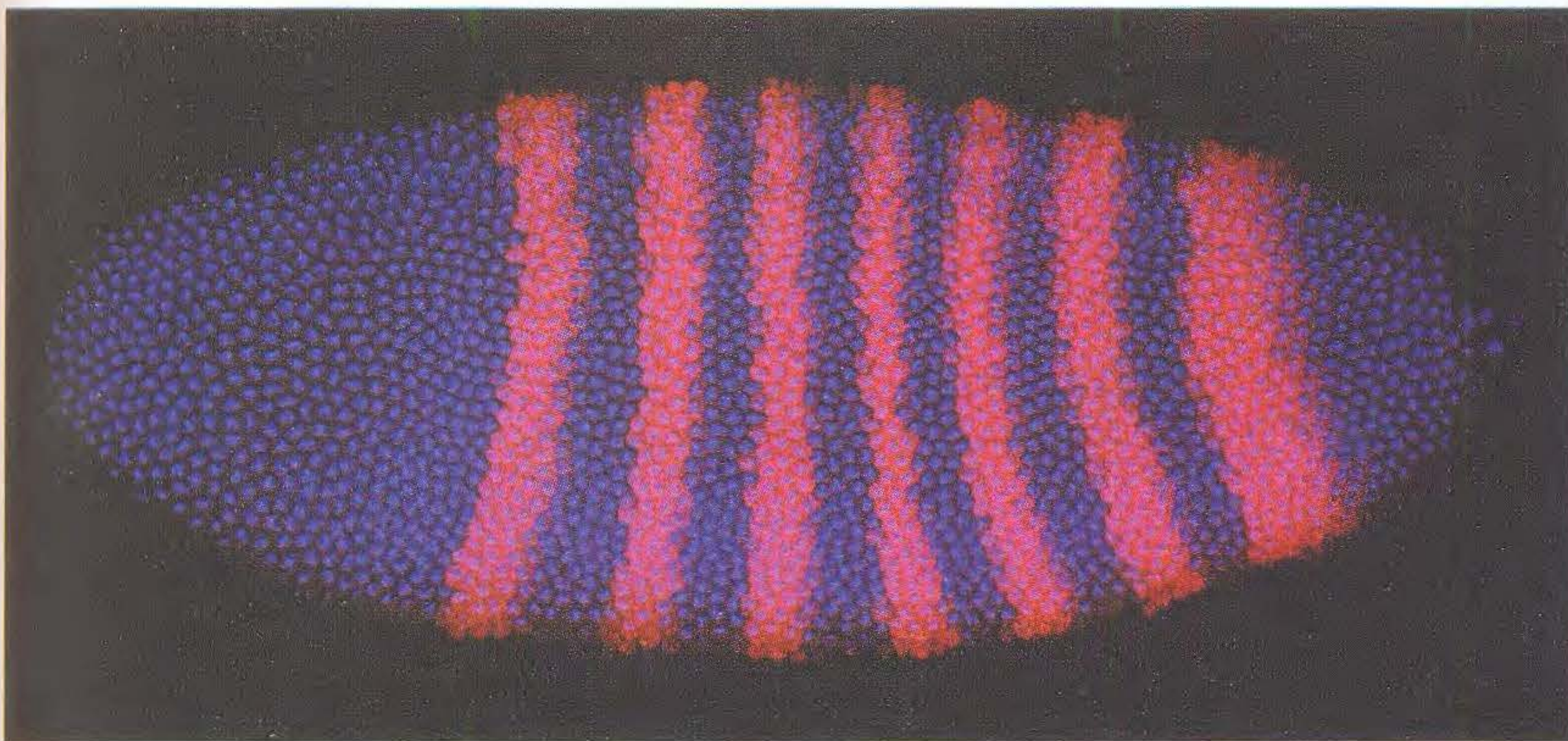
Você introduz todos os quatro desses clones simultaneamente em células de cultura de tecidos e, então, estimula aliquotas individuais dessas células com um dos três fatores. A análise de gel do RNA isolado dessas células deu os seguintes resultados. São mostrados os níveis de transcritos produzidos pelos genes introduzidos em resposta a vários tratamentos; a intensidade dessas bandas é proporcional à quantidade de transcritos feitos a partir de determinado clone. (A falta de aparecimento de uma banda parece indicar que o nível de transcritos é indetectável.)

Sem tratamento	FGF	Cort	EP	Clone
—	—		—	A
—	—		—	B
—	—	—	—	C
			—	D

Diagrama 2.

- a. Onde está o elemento de DNA que permite a ativação por FGF?
- b. Onde está o elemento de DNA que permite a repressão por Cort?
- c. Onde está o elemento de DNA que permite a indução por EP? Explique sua resposta.

O Controle Genético do Desenvolvimento



Expressão gênica de um desenvolvimento de embrião da mosca-das-frutas. As sete faixas vermelhas marcam as células que expressam o mRNA de um gene codificante de uma proteína reguladora que controla o número de segmentos no embrião de *Drosophila*. A regulação espacial da expressão gênica é central para o controle do desenvolvimento animal. [Foto de Dave Kosman, Ethan Bier, and Bill McGinnis.]

De todos os fenômenos em Biologia, poucos despertam mais interesse que a formação de um animal complexo a partir de uma única célula, o zigoto. Nessa transformação espetacular, várias forças organizam a massa de células em divisão em uma forma com uma cabeça e uma cauda distintas, vários apêndices e muitos órgãos. O grande geneticista Thomas Hunt Morgan não ficou imune a esse apelo estético:

Um zigoto transparente, à medida que se desenvolve, é um dos objetos mais fascinantes no mundo dos seres vivos. A contínua mudança de forma que ocorre de uma hora para outra nos desafia por sua simplicidade. Os padrões geométricos que se apresentam a cada mudança convidam a uma análise matemática.... Esse aparato é um apelo irresistível aos lados emocional e artístico de nossa natureza. [T.H. Morgan, *Experimental Embryology*. Columbia University Press, 1927.]

Mas, por toda a sua beleza e fascinação, os biólogos ficaram perplexos por muitas décadas com relação a como a forma biológica é gerada durante o desenvolvimento. Morgan também disse que "... se o mistério que envolve a Embriologia algum dia for compreendido, devemos ... recorrer a outros meios além da descrição do que ocorre".

A longa parada na Embriologia durou muito mais que os dias de Morgan nos anos de 1910 e 1920, mas foi resolvida por geneticistas que trabalharam muito na tradição da genética ao estilo de Morgan, e com seu modelo genético mais produtivo e favorito: a mosca-das-frutas *Drosophila melanogaster*.

Os principais catalisadores para compreender a formação dos animais foram as descobertas de monstros genéticos, moscas mutantes com marcan-

12

Perguntas Fundamentais

- Quais os genes que controlam o desenvolvimento, e como eles são identificados?
- Onde e quando esses genes estão ativos no curso do desenvolvimento?
- Como são controlados os genes reguladores de padrões?
- Como os genes reguladores de padrões afetam a forma animal?
- Diferentes grupos taxonômicos têm genes reguladores de padrões e processos em comum?

Tópicos

- 12.1 O enfoque genético do desenvolvimento
- 12.2 As ferramentas para o desenvolvimento de *Drosophila*
- 12.3 Definição de todas as ferramentas
- 12.4 Regulação espacial da expressão gênica no desenvolvimento
- 12.5 Regulação pós-transcricional da expressão gênica no desenvolvimento
- 12.6 Os vários papéis dos genes ferramentais individuais
- 12.7 Desenvolvimento e doença

tes alterações de estruturas corpóreas (Figura 12.1). Nos primeiros dias da genética de *Drosophila*, mutantes raros surgiram espontaneamente ou como subprodutos de outros experimentos com transformações espetaculares de partes do corpo. Em 1915, Calvin Bridges, então estudante de Morgan, isolou uma mosca com uma mutação que fazia com que os pequenos halteres das asas da mosca-das-frutas lembrassem as grandes asas anteriores. Ele chamou o mutante de *bithorax*. A transformação em mutantes *bithorax* é chamada de *homeótica* (do grego *homeos*, significando o mesmo ou similar) porque uma parte do corpo (a asa posterior) é transformada para se assemelhar à outra (à asa anterior), como mostrado na Figura 12.1b. Subseqüentemente, muitos outros mutantes homeóticos foram identificados em *Drosophila*, tais como o marcante mutante *Antennapedia*, no qual pernas desenvolvem-se no lugar de antenas (Figura 12.1c).

Os efeitos espetaculares dos mutantes homeóticos inspiraram o que iria tornar-se uma revolução na Embriologia, quando as ferramentas da Biologia Molecular tornaram-se disponíveis para compreender o que os genes homeóticos codificam e como eles exercem tamanha influência no desenvolvimento de partes inteiras do corpo. Surpreendentemente,

te, esses estranhos genes da mosca-das-frutas vieram a ser um passaporte para o estudo de todo o reino animal, pois contrapartes desses genes foram descobertas tendo papéis similares em quase todos os animais.

O estudo do desenvolvimento é uma disciplina grande e ainda crescente. Neste capítulo, enfocaremos alguns conceitos que ilustram a lógica do controle genético do desenvolvimento e como a informação para a construção de organismos complexos é codificada no genoma. O controle genético do desenvolvimento é fundamentalmente uma questão de regulação gênica no espaço tridimensional e com o tempo. Veremos que os princípios que governam o controle genético do desenvolvimento são conectados aos já apresentados nos Capítulos 10 e 11 para governar o controle fisiológico da expressão gênica em bactérias e eucariontes unicelulares.

► 12.1 O Enfoque Genético do Desenvolvimento

Por muitas décadas, o estudo do desenvolvimento embrionário envolveu a manipulação fisiológica dos embriões, células e tecidos. Foram estabelecidos vários conceitos principais sobre as propriedades de embriões em desenvolvimento por meio de experimentos nos quais uma parte de um embrião era transplantada para outra parte do embrião. Por exemplo, o transplante de uma parte de um embrião de anfíbio em desenvolvimento para outra parte em um embrião receptor foi mostrada induzindo o tecido circundante para formar um segundo eixo corpóreo completo (Figura 12.2a). Similarmente, o transplante da parte posterior de um membro de galinha em desenvolvimento para a anterior poderia induzir dedos extras, mas com polaridade revertida com relação aos dedos normais (Figura 12.2b). Essas regiões transplantadas do embrião de anfíbios e broto de embrião de galinha foram chamadas de *organizadores* devido à sua marcante habilidade em organizar o desenvolvimento dos tecidos vizinhos. As células nos organizadores foram postuladas produzindo *morfógenos*, moléculas que induzem várias respostas no tecido vizinho de um modo dependente da concentração.

Embora esses resultados experimentais fossem espetaculares e fascinantes, progressos posteriores na compreensão da natureza dos organizadores e morfógenos pararam após sua descoberta na primeira metade dos anos de 1900. Foi essencialmente impossível isolar moléculas responsáveis por essas atividades usando técnicas de separação bioquímica. As células embrionárias fazem milhares de substâncias, proteínas, glicoconjugados, hormônios e assim por diante. Um morfógeno pode ser qualquer uma dessas moléculas, mas estará presente em quantidades minúsculas — uma agulha no palheiro de produtos celulares.

O longo impasse em definir a Embriologia em termos moleculares foi quebrado por enfoques genéticos, principalmente o isolamento sistemático dos mutantes com pequenos defeitos no desenvolvimento e subseqüente caracterização e estudo dos produtos gênicos que eles codificam. O enfoque genético para estudar o desenvolvimento apresenta muitas vantagens em relação à alternativa das estratégias bioquí-

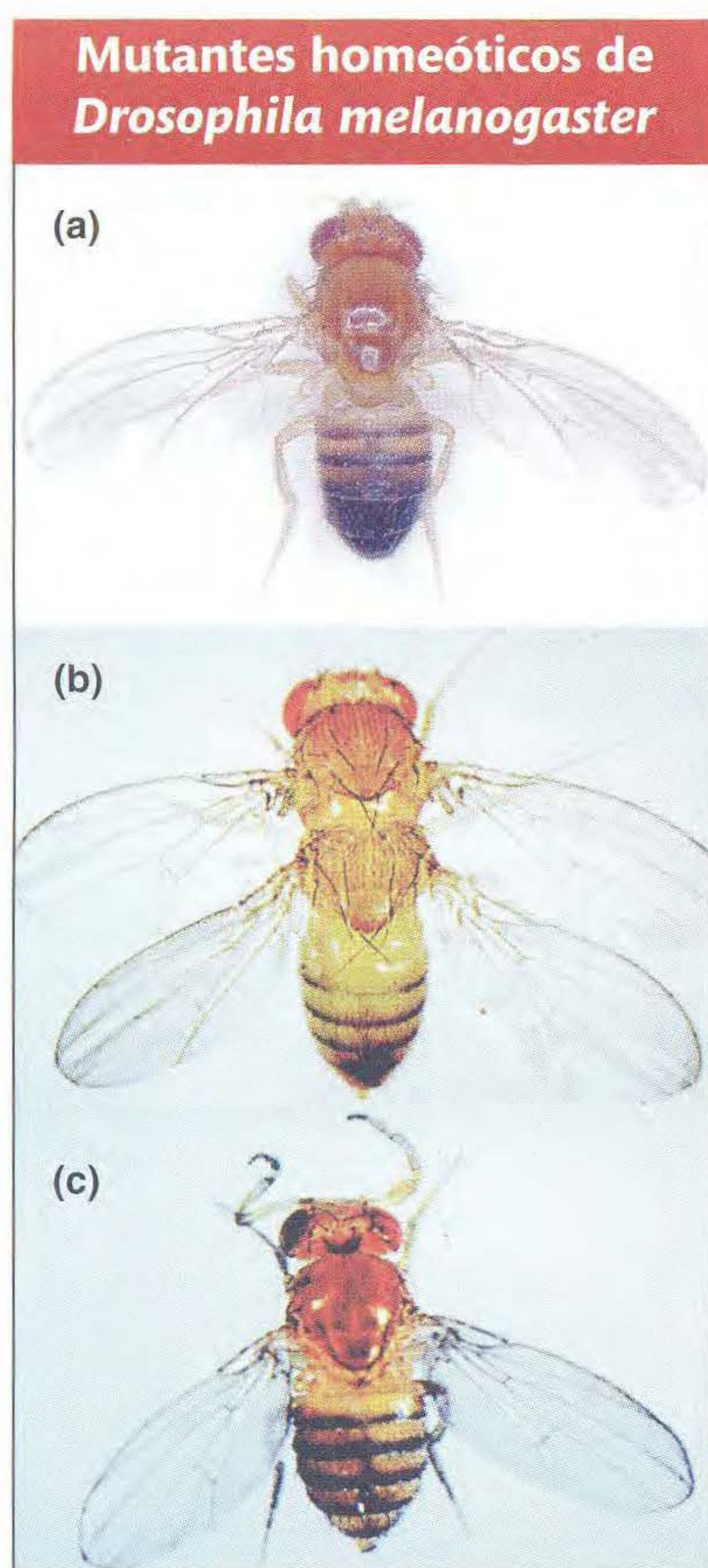


FIG. 12.1 Nos mutantes homeóticos, a identidade de uma estrutura corpórea foi alterada para outra. (a) Mosca normal com um par de asas anteriores no segundo segmento torácico e um par de pequenas asas posteriores no terceiro segmento torácico. (b) Mutante triplo para três mutações no gene *Ultrabithorax*. A função de *Ubx* é perdida no tórax posterior, o que causa o desenvolvimento de asas anteriores no lugar das asas posteriores. (c) O mutante *Antennapedia* no qual as antenas são transformadas em pernas. [Cortesia de Sean Carroll.]

Organizadores em embriões animais

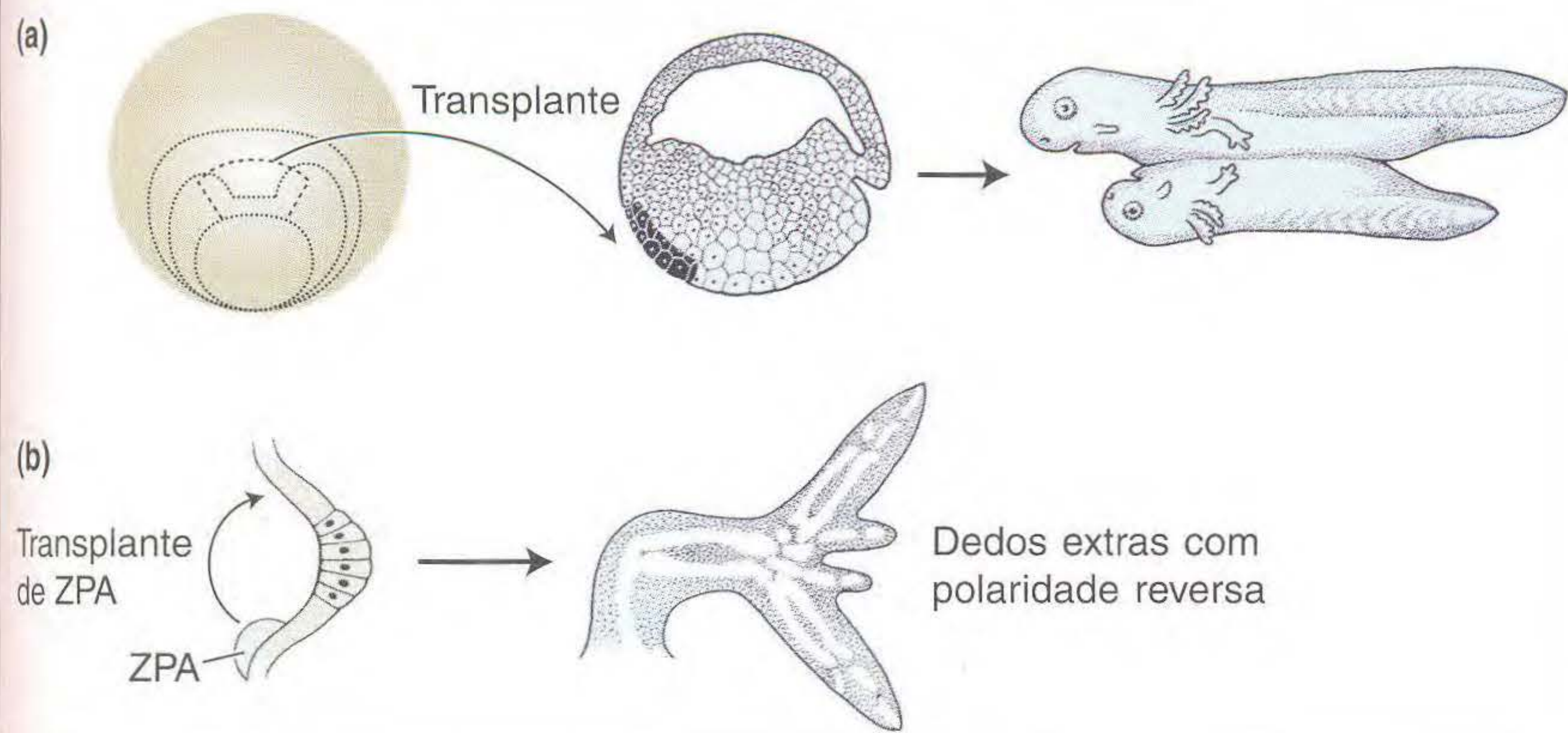


FIG. 12.2 Experimentos de transplante tiveram um papel central no início da Embriologia, e demonstraram a atividade de organização de longo alcance dos tecidos embrionários. (a) O organizador Spemann. O "lábio" do blastômero dorsal de um anfíbio inicial pode induzir um segundo eixo embrionário e o embrião quando transplantado para a região ventral de um embrião receptor. (b) No broto do membro vertebral do desenvolvimento de galinha, a zona de polarização de atividade (ZPA) organiza padrões ao longo do eixo ântero-posterior. O transplante de ZPA de uma posição posterior para uma anterior induz dedos extras com polaridade reversa.

micas. Primeiro, o geneticista não precisa fazer nenhuma suposição sobre o número ou natureza das moléculas necessárias para um processo. Segundo, a (limitada) quantidade de um produto gênico não é um impedimento: todos os genes podem ser mutados independentemente da quantidade do produto feito pelo gene. Terceiro, o enfoque genético pode descobrir fenômenos para os quais não há testes bioquímicos ou outros.

Do ponto de vista genético, existem quatro perguntas principais envolvendo o número, identidade e função dos genes que tomam parte no desenvolvimento:

1. Que genes são importantes no desenvolvimento?
2. Onde, no desenvolvimento animal, e em que épocas esses genes estão ativos?
3. Como a expressão dos genes desenvolvimentais é regulada?
4. Por que mecanismos moleculares os produtos gênicos afetam o desenvolvimento?

Para abordar essas perguntas, têm que ser criadas estratégias para identificar, catalogar e analisar genes que controlam o desenvolvimento. Uma das primeiras considerações na análise genética do desenvolvimento animal foi que animal estudar. Dos milhões de espécies vivas, qual a mais promissora? A mosca-das-frutas *Drosophila melanogaster* surgiu como o modelo genético líder do desenvolvimento animal porque é mais fácil de criar, tem ciclo reprodutivo rápido, e a citogenética e décadas de análise genética clássica (incluindo o isolamento de muitos mutantes marcantes) deram importantes vantagens experimentais (veja o boxe Organismo-modelo *Drosophila*, adiante). O verme nematódeo *Caenorhabditis elegans* também apresenta muitas características atraentes, mais particularmente sua construção simples e linhagens celulares bem estudadas. Entre os vertebrados, o desenvolvimento das técnicas de perturbação de genes-alvo abriu o camundongo de laboratório para estudos genéticos mais sistemáticos, e o peixe zebra *Danio rerio* recentemente se tornou um modelo favorito devido à transparência do embrião e aos avanços em seu estudo genético.

Pela análise genética sistemática e dirigida, bem como estudos genômicos comparativos, grande parte das ferramen-

tas genéticas para o desenvolvimento de várias espécies diferentes foram definidas. Primeiro enfocaremos as ferramentas de *Drosophila melanogaster*, pois sua identificação foi uma fonte de importantes *insights* no controle genético do desenvolvimento; sua descoberta catalisou a identificação das ferramentas para outros animais, incluindo humanos.

► 12.2 As Ferramentas para o Desenvolvimento de *Drosophila*

Os genomas animais tipicamente contêm cerca de 12.000 a 25.000 genes. Muitos desses genes codificam proteínas que funcionam em processos essenciais em todas as células do corpo (por exemplo, no metabolismo celular ou na biossíntese de macromoléculas). Tais genes em geral são chamados de **genes de manutenção**. Outros genes codificam proteínas que efetuam tarefas especializadas de vários sistemas orgânicos, tecidos e células do corpo, tais como as proteínas globinas no transporte de oxigênio ou proteínas de anticorpos que medeiam a imunidade. Aqui, estamos interessados em um conjunto diferente de genes, os envolvidos com a estrutura dos órgãos e tecidos, e a especificação dos tipos celulares, as ferramentas que determinam o plano corpóreo geral e o número, identidade e padrão das partes do corpo.

Os genes ferramentais (*toolkit*) da mosca-das-frutas foram em geral identificados pelas monstruosidades ou catástrofes que surgem quando eles estão mutados. As mutações de genes ferramentais de duas fontes produziram a maioria de nossos conhecimentos. A primeira fonte consiste em mutações espontâneas que surgem nas populações de laboratório. A segunda fonte compreende mutações induzidas aleatoriamente pelo tratamento com mutágenos (tais como substâncias químicas ou radiação) que aumentam muito a frequência de genes danificados pelo genoma. Refinamentos elegantes deste último enfoque permitiram pesquisas sistemáticas de mutantes para mutações que identificaram muitos membros das ferramentas genéticas da mosca. Os membros dessas ferramentas constituem apenas uma pequena fração, talvez várias centenas de genes, dos aproximadamente 14.000 genes do genoma da mosca.



Organismo-modelo *Drosophila*

Análise Mutacional do Início do Desenvolvimento de *Drosophila*

Os *insights* iniciais sobre o controle genético da formação de padrão emergiu de estudos da mosca-das-frutas *Drosophila melanogaster*. O desenvolvimento de *Drosophila* provou ser uma mina de ouro para os pesquisadores porque os problemas desenvolvimentais podem ser abordados com o uso simultâneo de técnicas genéticas e moleculares.

O embrião de *Drosophila* foi especialmente importante para a compreensão da formação do plano corpóreo animal básico. Um importante motivo é que uma anomalia no plano corpóreo de um mutante é facilmente identificada no exoesqueleto larvar no embrião de *Drosophila*. O exoesqueleto larvar é uma estrutura não-celular, feita de um polímero de polissacarídeo, chamado quitina, que é produzido como uma secreção das células epidérmicas do embrião. Cada estrutura do exoesqueleto é formada de células epidérmicas ou células imediatamente subjacentes a essa estrutura. Com seu padrão complexo de cerdas, indentações e outras estruturas, o exoesqueleto fornece vários marcos para servir como indicadores dos destinos determinados para muitas células epidérmicas. Em particular, existem muitas estruturas anatômicas distintas ao longo dos eixos ântero-posterior (A-P) e dorsoventral (D-V). Além disso, como todos os nutrientes necessários ao desenvolvimento para o estágio larval estão pré-embalados no zigoto, os embriões mutantes nos quais os destinos celulares A-P ou D-V estão marcadamente alterados podem se desenvolver até o final da embriogênese e produzir uma larva mutante em cerca de 1 dia (veja diagrama). O exoesqueleto de tal larva mutante reflete os destinos mutantes atribuídos a subgrupos

de células epidérmicas, e podem, assim, identificar genes merecedores de uma análise detalhada.

O desenvolvimento do padrão corpóreo adulto de *Drosophila* leva um pouco mais de uma semana (veja ilustração). Pequenas populações de células postas de lado durante a embriogênese proliferam durante três estágios larvares (instars) e diferenciam-se no estágio de pupa até as estruturas adultas. Essas células postas de lado incluem os *discos imaginais*, que são regiões em forma de disco que dão origem a apêndices específicos e tecidos em cada segmento, tais como pernas, asas, olhos e discos das antenas. Os discos imaginais são fáceis de remover para análise da expressão gênica (veja Figura 12.7).

Os genes que contribuem para o plano corpóreo de *Drosophila* podem ser clonados e caracterizados em âmbito molecular com facilidade. A análise dos genes clonados em geral dá valiosas informações sobre a função do produto proteico, geralmente identificando correlatos próximos em sequência de aminoácidos do polipeptídeo codificado através de comparações com todas as sequências de proteínas estocadas em bancos de dados públicos. Além disso, podemos investigar os padrões espaciais e temporais de expressão de (1) um mRNA, usando sequências de DNA unifilamentares histoquimicamente marcadas complementares ao mRNA para fazer hibridização *in situ*, ou (2) uma proteína, usando anticorpos histoquimicamente marcados que se ligam especificamente a essa proteína.

Uso do Conhecimento de um Organismo-modelo para Rastrear o Desenvolvimento de Descobertas Gênicas em Outros

Com a descoberta de que existem vários genes homeobox dentro do genoma de *Drosophila*, as similaridades entre

As ferramentas genéticas para o desenvolvimento animal são uma pequena fração de todos os genes. Apenas um pequeno subgrupo do complemento inteiro de genes no genoma afeta o desenvolvimento de modos discretos.

Classificação dos genes pela função desenvolvimental

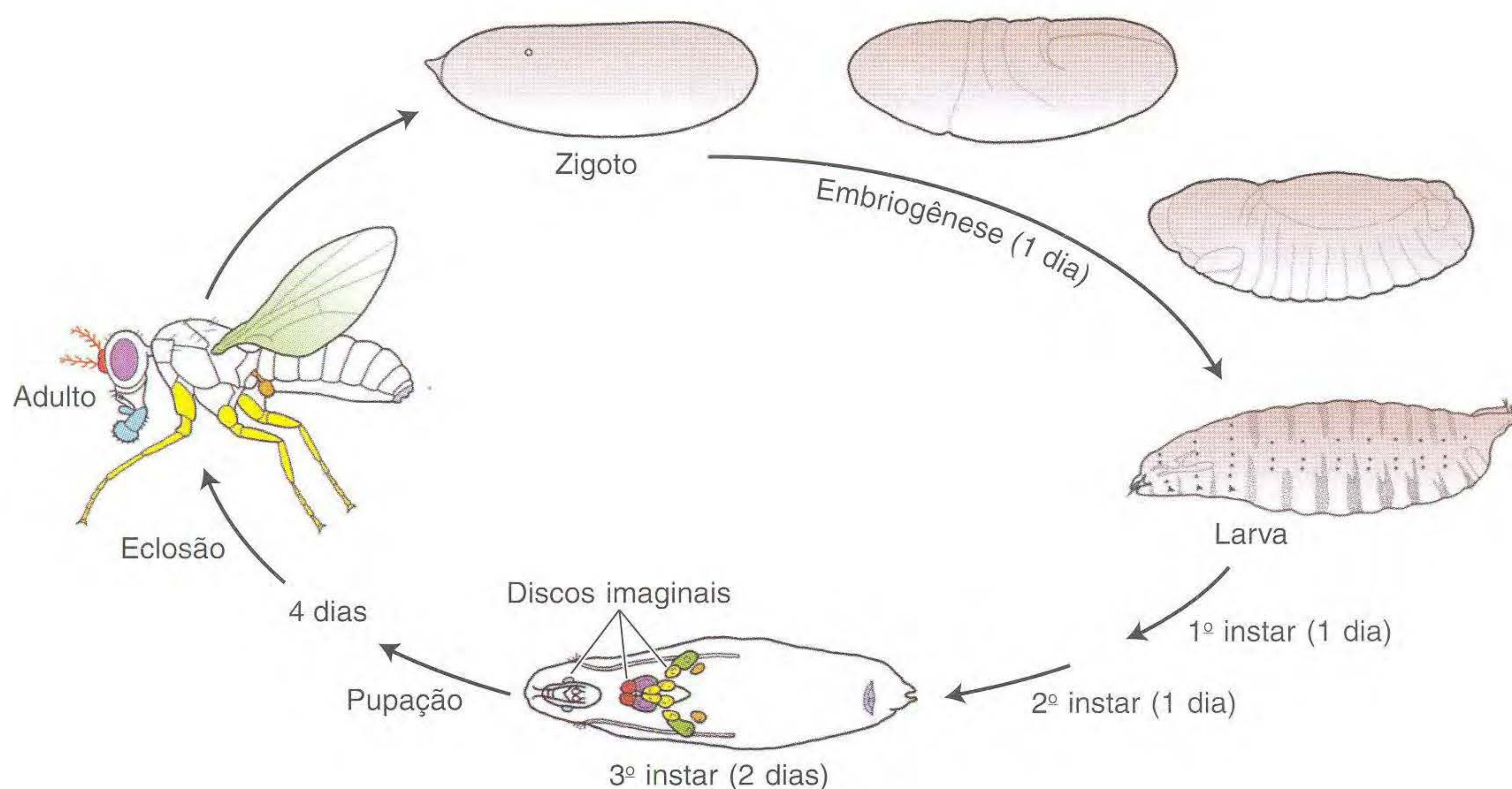
Uma das primeiras tarefas após a execução de uma triagem genética para mutações é separar aquelas de interesse. Muitas mutações são letais quando hemi- ou homozigotas porque as células não podem sobreviver sem os produtos afetados por essas mutações. As mutações mais interessantes são as que causam algum pequeno defeito no padrão corpóreo embrionário e/ou do adulto. Demonstrou-se útil agrupar os genes afetados por mutações em várias categorias com base na natureza de seus fenótipos mutantes. Muitos genes ferramentais podem ser classificados de acordo com sua função em controlar a identidade de partes do corpo (por exemplo,

de segmentos diferentes ou apêndices), a formação de partes do corpo (por exemplo, de órgãos ou apêndices), o número de partes do corpo, a formação de tipos de células e a organização dos eixos corpóreos primários (o ântero-posterior, ou A-P, e o dorsoventral, ou eixo D-V).

Começaremos nosso inventário das ferramentas de *Drosophila* examinando os genes que controlam a identidade de segmentos e apêndices. Fazemos isso com fins históricos e conceituais. Os genes que controlam a identidade segmental e de apêndices estão entre os primeiros genes ferramentais identificados. As descobertas subsequentes sobre sua natureza foram fontes de profundos *insights* sobre como seus produtos funcionam, mas também o conteúdo e funcionamento das ferramentas da maioria dos animais. Além disso, seus espetaculares fenótipos mutantes indicam que eles estão entre os genes globalmente mais atuantes que afetam a forma animal. Aprender sobre esses genes deve estimular nosso apetite por aprender mais sobre todas as ferramentas que controlam o desenvolvimento da forma animal.

as seqüências de DNA desses genes podem ser exploradas em caçadas a outros membros da família do gene homeótico. Essas caçadas dependem da complementariedade dos pares de bases do DNA. Para esse fim, as hibridizações de DNA foram feitas sob *condições de moderado rigor*, nas quais pode haver alguns erros de pareamentos de bases entre os padrões de hibridização sem perturbar o apropriado pareamento de pontes de hidrogênio de pares de bases vizinhas. Algumas dessas caçadas foram feitas no próprio genoma de

Drosophila, à procura de membros mais familiares. Outras procuravam genes homeobox em outros animais, por meio de *zoo blots* (transferências de Southern de DNA digerido por enzima de restrição de animais diferentes), usando DNA homeobox como sonda. Esse enfoque levou à descoberta de seqüências homeobox homólogas em muitos animais diferentes, incluindo humanos e camundongos. (De fato, é um enfoque muito poderoso para “pescar” correlatos de quase todos os genes em seu organismo favorito.)



Visão geral do desenvolvimento de *Drosophila*. A larva forma-se em 1 dia e, então, sofre vários estágios de crescimento durante os quais os discos imaginários e outros precursores das estruturas adultas proliferam. Essas estruturas diferenciam-se durante a pupação, e a mosca adulta eclode e começa novamente o ciclo.

Genes homeóticos e identidade segmentar

Entre as anomalias mais fascinantes a serem descritas nos animais estão aquelas nas quais uma parte normal do corpo é substituída por outra. Tais transformações homeóticas foram observadas em muitas espécies na natureza, incluindo tentredém, na qual um perna forma-se no lugar de uma antena, e sapos, nos quais as vértebras torácicas formam-se no lugar de uma vértebra cervical (Figura 12.3). Enquanto apenas um membro de um par bilateral de estruturas é comumente alterado em muitas variantes de ocorrência natural, ambos os membros de um par bilateral de estruturas são alterados em mutantes homeóticos das moscas-das-frutas. Tais mutantes homeóticos reproduzem-se de geração a geração.

O fascínio científico com mutantes homeóticos é baseado em três propriedades. Primeira, é curioso que uma única mutação gênica possa alterar uma via de desenvolvimento tão intensamente. Segundo, é marcante que a estrutura formada no mutante seja um similar bem desenvolvido de outra

parte do corpo. Terceiro, é importante notar que as mutações homeóticas transformam a identidade de **estruturas serialmente reiteradas**. Os insetos e muitos corpos de animais são feitos de partes repetidas de estruturas similares, como blocos estruturais, dispostos em uma série. As asas anteriores e posteriores, os segmentos e as antenas, pernas e partes da boca dos insetos são partes do corpo serialmente reiteradas. As mutações homeóticas transformam identidades dentro desses grupos.

Uma mutação pode causar uma perda de função de um gene homeótico onde o gene normalmente atua, ou pode causar um ganho de função homeótica onde o gene homeótico normalmente não atua. Por exemplo, o gene *Ultrabithorax* (*Ubx*) atua no desenvolvimento da asa posterior para desenvolver o desenvolvimento da asa posterior e reprimir o desenvolvimento da asa anterior. A perda de função das mutações em *Ubx* transforma a asa posterior em uma anterior. As mutações dominantes de ganho de função em *Ubx* transformam a asa anterior em uma asa posterior. Similar-

Em uma transformação homeótica, uma parte do corpo é substituída por outra

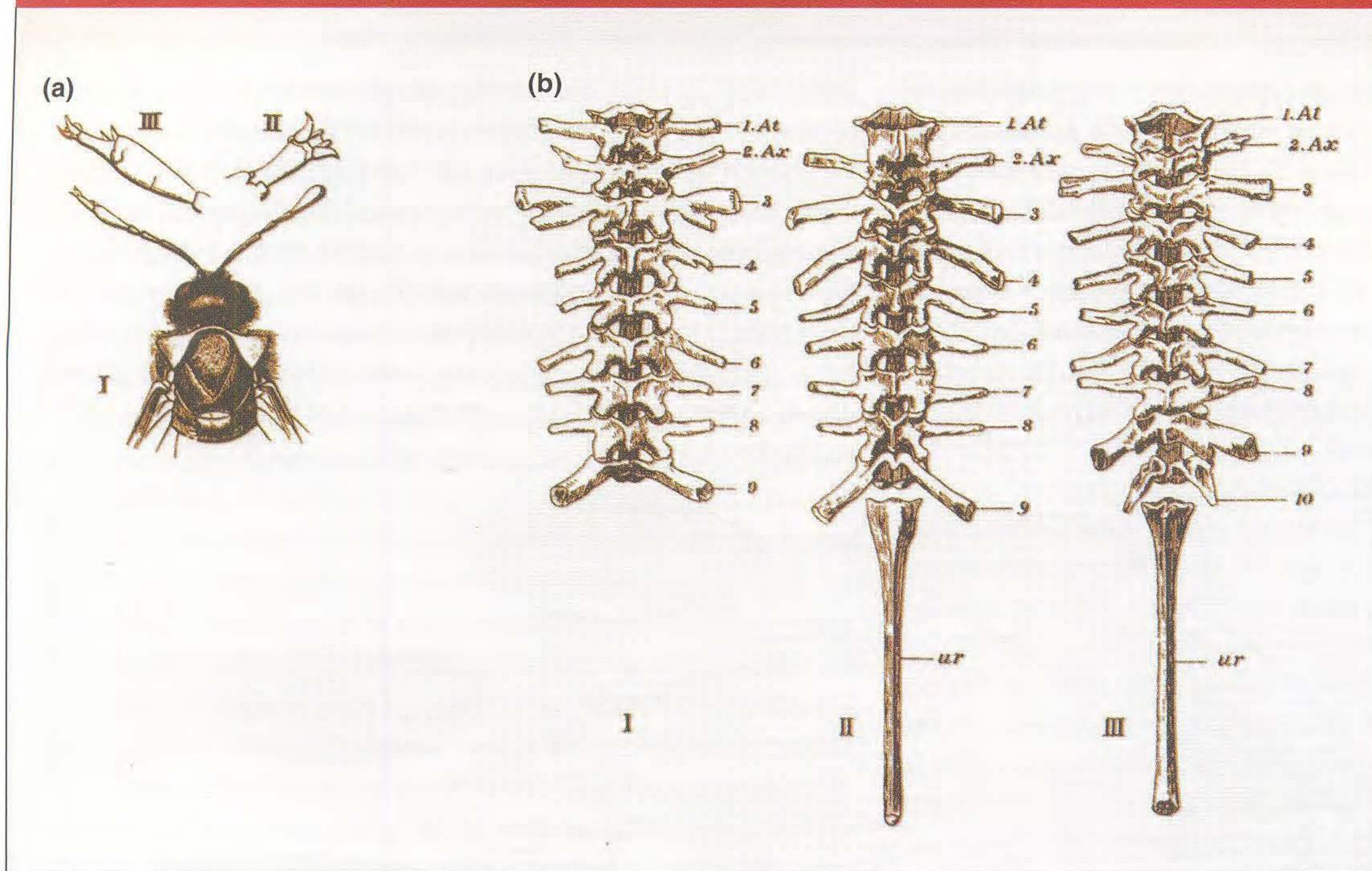


FIG. 12.3 Um desenho do final do século XIX de um dos primeiros estudos de transformações homeóticas na natureza. (a) Homeose em tentredém, com a antena esquerda transformada em uma perna. (b) Homeose em um sapo. O animal do meio é normal. O da esquerda tem estruturas extras crescendo fora do topo da coluna vertebral. A amostra da direita tem um conjunto extra de vértebras. [De W. Bateson, *Material for the Study of Variation*. Macmillan. 1894.]

mente, as transformações antena para perna dos mutantes *Antennapedia* (*Antp*) são causadas pelo ganho dominante de função *Antp* na antena. Além dessas transformações na identidade dos apêndices, as mutações homeóticas podem transformar a identidade do segmento, fazendo com que um segmento corpóreo do adulto ou larva se assemelhe a outro.

Embora os genes homeóticos fossem primeiro identificados por mutações espontâneas afetando moscas adultas, eles são necessários durante a maior parte do desenvolvimento de uma mosca. As pesquisas sistemáticas de genes homeóticos levaram à identificação de oito loci, hoje chamados de **genes Hox**, que afetam a identidade de segmentos e seus apêndices associados em *Drosophila*. Geralmente, a perda completa de qualquer função do gene *Hox* é letal no início do desenvolvimento. As mutações dominantes que transformam adultos são viáveis nos heterozigotos porque o alelo tipo selvagem dá a função gênica normal para o animal em desenvolvimento.

Organização e expressão dos genes Hox

Uma característica curiosa dos genes *Hox* é que eles são agrupados em dois **complexos gênicos** que estão situados no terceiro cromossomo de *Drosophila*. O complexo *Bithorax* contém três genes *Hox*, e o complexo *Antennapedia* contém cinco genes *Hox*. Além disso, a ordem dos genes nos complexos e no cromossomo corresponde à ordem das regiões do corpo, da cabeça até a cauda, que são influenciadas por cada gene *Hox* (Figura 12.4).

A relação entre a estrutura dos complexos gene *Hox* e os fenótipos dos mutantes gene *Hox* foi iluminada pela carac-

terização molecular dos genes. A clonagem molecular das seqüências que englobam cada locus *Hox* fornece um meio para analisar onde, no desenvolvimento animal, cada gene é expresso. Esses aspectos espaciais da expressão gênica e regulação gênica são cruciais para compreender a lógica do controle genético do desenvolvimento. Com relação aos genes

Genes Hox regulam a identidade das partes do corpo

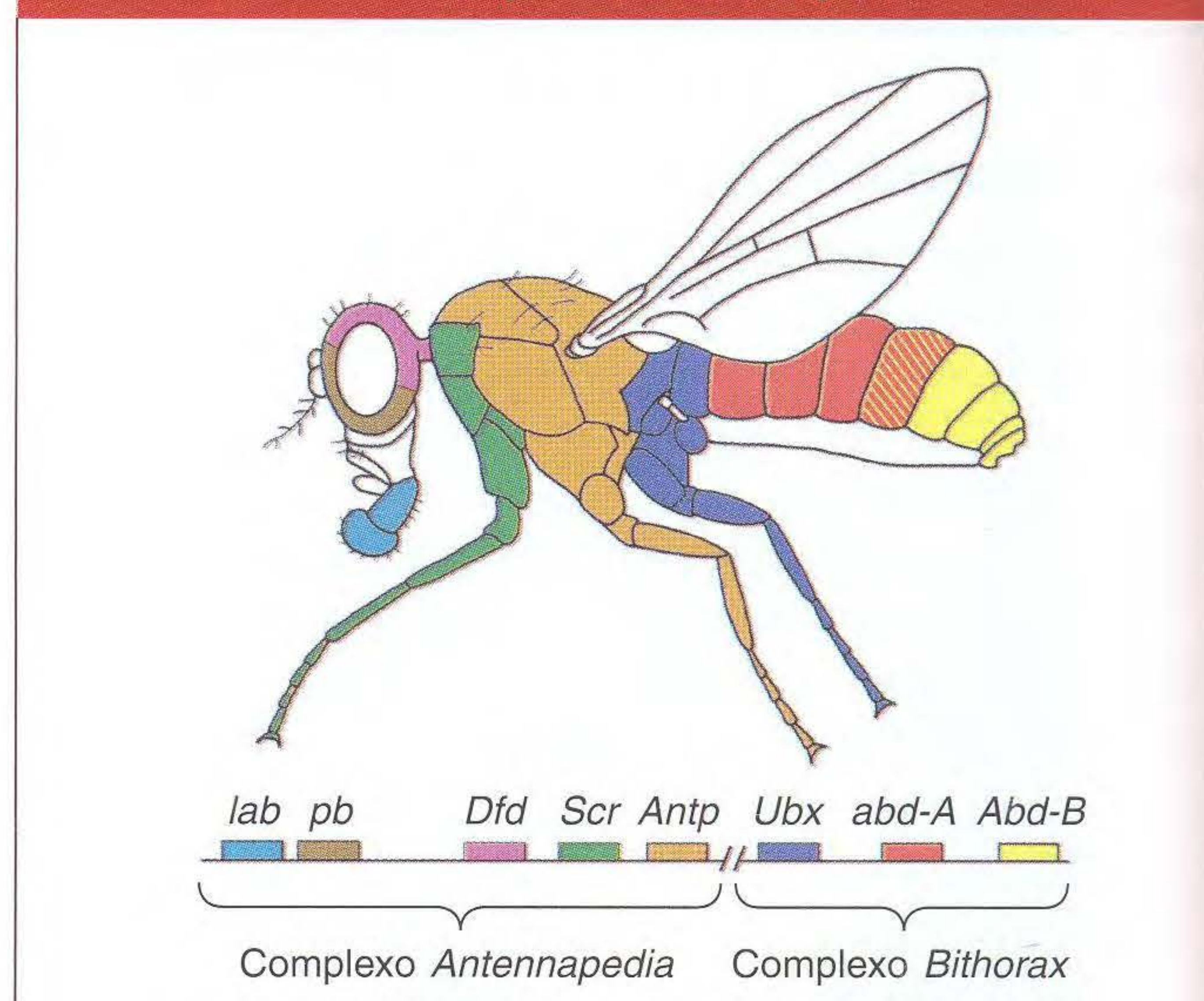


FIG. 12.4 Os genes *Hox* de *Drosophila*. Oito genes *Hox* regulam a identidade das regiões dentro do adulto. As cores identificam os segmentos e estruturas que são afetados pelas mutações nos vários genes *Hox*. [De S. B. Carroll, J. K. Grenier, and S. D. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*, 2nd ed. Blackwell, 2005.]

Métodos para visualização da expressão gênica nos animais em desenvolvimento

Hibridização *in situ* para visualização de mRNA transcritos

Transcreve para obter uma sonda de RNA unifilamentar com nucleotídeos modificados complementares à sequência de mRNA.

clone de cDNA de um gene

Imunolocalização da expressão da proteína

Proteína expressa em bactérias.

Injetar proteína no vertebrado hospedeiro.

Extrair anticorpos à proteína (IgG).

Embriões fixados ou tecido dissecado

Incubar embriões com sonda de RNA (sonda hibridiza com mRNA).

Eliminar sonda não-ligada.

Adicionar anticorpo conjugado a enzima a um nucleotídeo específico modificado.

Eliminar anticorpo não-ligado.

Adicionar substrato da enzima.

Visualizar expressão de mRNA ao microscópio óptico.

Incubar embriões com anticorpos (anticorpo liga-se à proteína).

Eliminar anticorpo não-ligado.

Adicionar anticorpo conjugado a fluorocromo à IgG da espécie hospedeira.

Eliminar anticorpo não-ligado.

Visualizar expressão de proteína ao microscópio de fluorescência.



FIG. 12.5 As duas principais tecnologias para visualizar onde um gene é transcrito ou onde a proteína que ele codifica é expressa são (*esquerda*) hibridização *in situ* de sonda complementar de RNA ao mRNA e (*direita*) imunolocalização da expressão da proteína. Os procedimentos para cada método são descritos. Os padrões de expressão podem ser visualizados como o produto de uma reação enzimática ou de um substrato cromogênico ou com compostos marcados fluorescentemente.

Hox e outros genes ferramentais, o desenvolvimento da tecnologia que permitiu a visualização da expressão do gene e proteína foi crucial para a compreensão da relação entre a organização gênica, funcionamento gênico e fenótipos mutantes.

Duas tecnologias principais para visualização da expressão gênica em embriões e outros tecidos são: (1) a expressão dos RNA transcritos visualizada por hibridização *in situ* e (2) a expressão das proteínas *Hox* visualizadas por métodos imunológicos. Cada tecnologia depende do isolamento de clones de cDNA representando o mRNA transcrito final e as proteínas (Figura 12.5).

No embrião em desenvolvimento, os genes *Hox* são expressos em domínios espacialmente restritos, às vezes superpostos dentro do embrião (Figura 12.6). Os genes são também

expressos nos tecidos da larva e pupa que darão origem a partes do corpo do adulto.

Os padrões de expressão do gene *Hox* (e outros genes ferramentais) em geral estão correlacionados com as regiões do animal afetadas pelas mutações gênicas. Por exemplo, o sombreado azul na Figura 12.6 indica onde o gene *Ubx* é expresso. Esse gene *Hox* é expresso no segmento torácico posterior e na maioria dos segmentos abdominais do embrião. O desenvolvimento desses segmentos é alterado nos mutantes *Ubx*. *Ubx* também é expresso na asa posterior, mas não na asa anterior em desenvolvimento (Figura 12.7), como seria esperado sabendo que *Ubx* promove o desenvolvimento da asa posterior e reprime o desenvolvimento da asa anterior nesse apêndice.

Os genes *Hox* são expressos em domínios espacialmente restritos

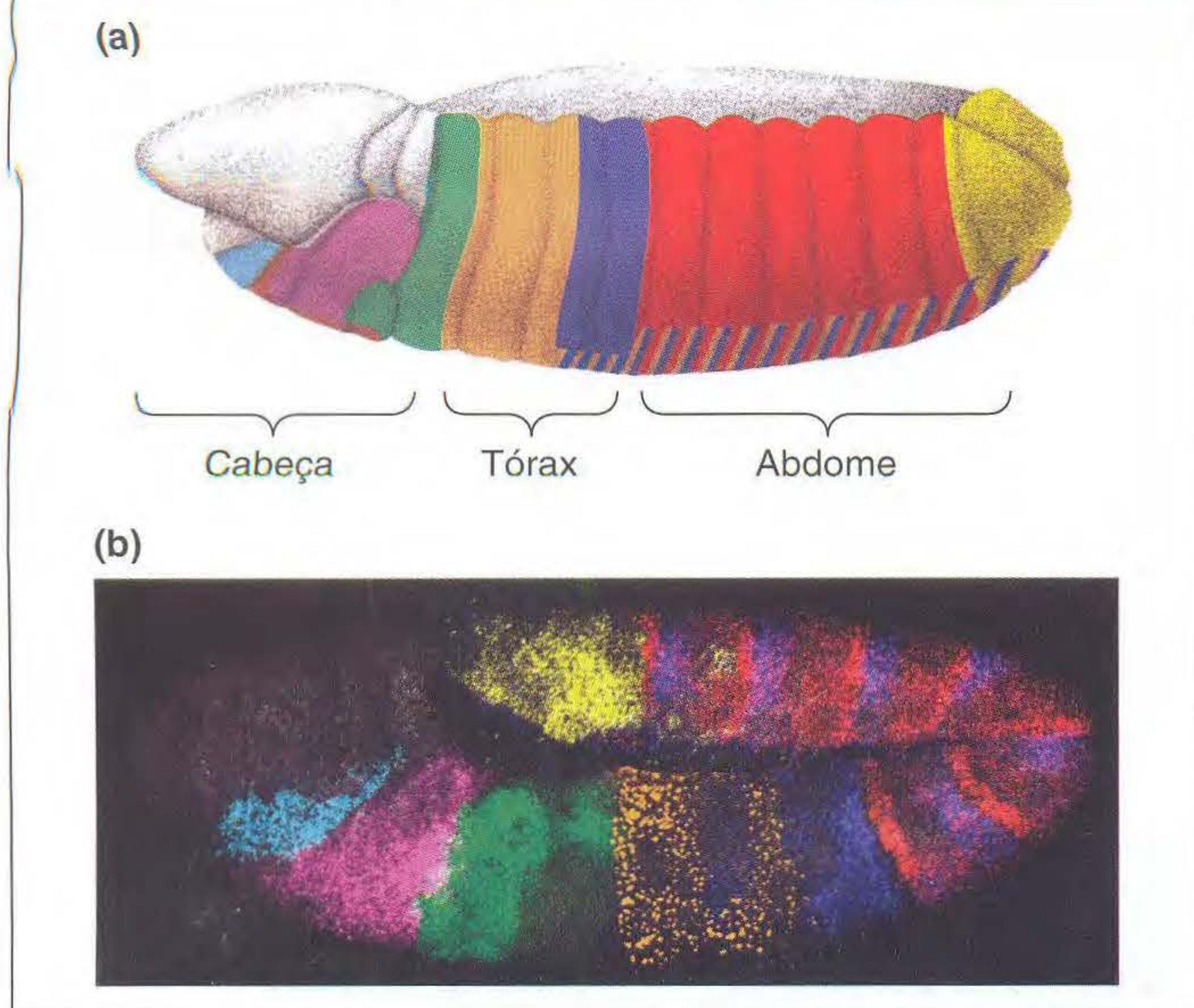


FIG. 12.6 Expressão dos genes *Hox* no embrião de *Drosophila*. (a) Representação esquemática do embrião de *Drosophila* mostrando regiões onde os oito genes *Hox* individuais são expressos. (b) Imagem da expressão dos sete genes *Hox* visualizados por hibridização *in situ*. As cores indicam a expressão de *labial* (turquesa), *Deformed* (púrpura), *Sex combs reduced* (verde), *Antennapedia* (laranja), *Ultrabithorax* (azul), *Abdominal-A* (vermelho) e *Abdominal-B* (amarelo). O embrião está dobrado, de modo que a extremidade posterior (amarela) aparece perto do centro. [(b) Fotomicrografia de Dave Kosman, Ethan Bier, and Bill McGinnis.]

A expressão espacial dos genes ferramentais geralmente está proximamente correlacionada a regiões do animal afetadas por mutações gênicas.

É crucial para distinguir o papel dos genes *Hox* em determinar a *identidade* de uma estrutura da que governa a sua *formação*. Na ausência de função de todos os genes *Hox*, os segmentos formam-se, mas têm a mesma identidade; os

membros também podem formar-se, mas têm a identidade de antenas anteriores. Outros genes controlam a formação de segmentos, membros e asas, e serão descritos mais adiante. Primeiro, temos que compreender como os genes *Hox* exercem seus efeitos marcantes no desenvolvimento da mosca.

O homeobox

Como os genes *Hox* têm grandes efeitos nas identidades de segmentos inteiros e outras estruturas corpóreas, a natureza e função das proteínas que eles codificam são de interesse especial. Edward Lewis, um pioneiro no estudo de genes homeóticos, notou que o aglomerado do complexo de genes *Bithorax* sugeriu que vários loci tinham surgido por duplicação em tandem de um gene ancestral. Essa idéia levou os pesquisadores a procurar similaridades nas seqüências de DNA dos genes *Hox*. Eles encontraram que todos os oito genes *Hox* dos dois complexos eram suficientemente similares para se hibridizarem. Essa hibridização foi vista como sendo devida a uma curta região da seqüência de cada gene, 180 pb de tamanho. Esse trecho de similaridade da seqüência de DNA, devido à sua presença em genes homeóticos, foi chamado de **homeobox**. O homeobox codifica um domínio de proteína, o **homeodomínio**, contendo 60 aminoácidos. A seqüência de aminoácidos do homeodomínio é muito similar entre as proteínas *Hox* (Figura 12.8).

Embora a descoberta de um motivo proteico comum em cada uma das proteínas *Hox* fosse muito entusiasmante, análises posteriores da estrutura do homeodomínio revelaram que ela forma um motivo hélice-giro-hélice comum ao repressor Lac, repressor λ , Cro e proteínas reguladoras $\alpha 2$ e $\alpha 1$ dos loci de tipos reprodutivos de leveduras. Essa similaridade sugeriu imediatamente (e subsequentelemente mostrou) que as proteínas *Hox* são proteínas de ligação ao DNA específicas de seqüências, e que elas exercem seus efeitos controlando a expressão de genes dentro de segmentos de desenvolvimento e apêndices. Assim, os produtos desses genes marcantes funcionam através de princípios que já são conhecidos dos Capítulos 10 e 11, ligando-se a elementos regulatórios de outros genes para ativar ou reprimir sua expressão. Veremos que isso também é

Genes *Hox* são expressos em estruturas afetadas por mutações em gene *Hox*

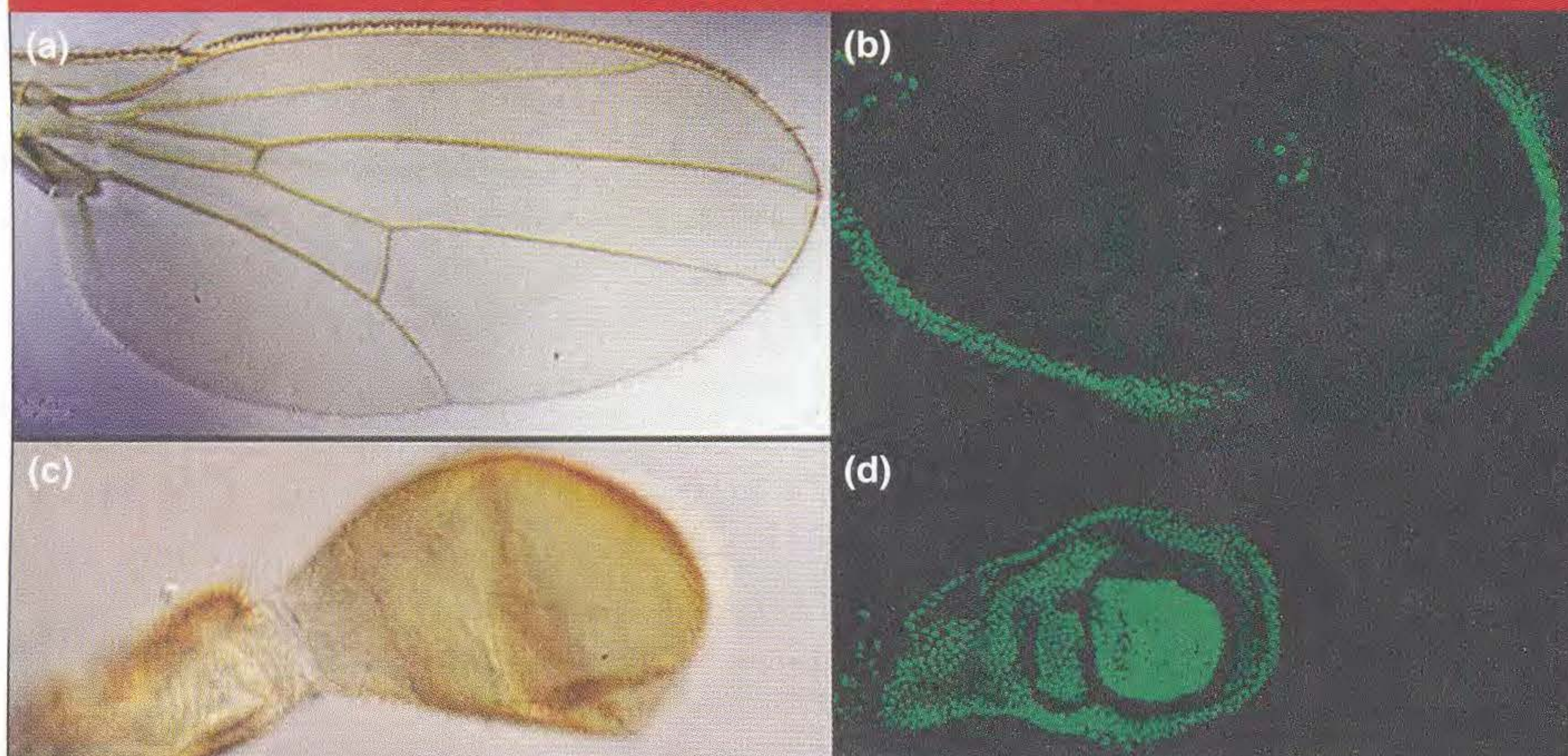


FIG. 12.7 Um exemplo da expressão do gene *Hox*. (a) A asa adulta de *D. melanogaster*. (b) A proteína Ubx não é expressa em células do disco imaginal em desenvolvimento que irá formar a asa anterior. As células enriquecidas de proteínas *Hox* são coradas de verde; nesta imagem, as células coradas em verde são células que não formam a asa. (c) O haltere da asa adulta posterior. (d) A proteína Ubx é expressa em altos níveis em todas as células do disco imaginal da asa posterior. [Fotos de Scott Weatherbee.]

As proteínas Hox têm uma sequência em comum

<i>lab</i>	NNSGRTNFTNKQLTELEKEFHFNRYLTRARRIEIANLQNETQVKIWFQNRRMKQKKRV
<i>pb</i>	PRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQT
<i>Dfd</i>	PKRQRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKKDN
<i>Scr</i>	TKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEH
<i>Antp</i>	RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN
<i>Ubx</i>	RRRGRQTYTRYQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQIKIWFQNRRMKLKKET
<i>abd-A</i>	RRRGRQTYTRFQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQIKIWFQNRRMKLKKEL
<i>abd-B</i>	VRKKRKPYSKFQTLELEKEFLFNAYVSKQKRWELARNLQALTERQVKIWFQNRRMKKNKNS
Seqüência de consenso	-RRGRT-YTR-QTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMK-KKE-
	Hélice 1 Hélice 2 Hélice 3

FIG. 12.8 Seqüências de homeodomínios da mosca. Todos os oito genes *Hox* de *Drosophila* codificam proteínas contendo um domínio altamente conservado de 60 aminoácidos, o homeodomínio, composto de três hélices α . As hélices 2 e 3 formam um motivo hélice-giro-hélice semelhante ao repressor Lac, Cro e outras proteínas de ligação ao DNA. As unidades comuns aos genes *Hox* são sombreadas em amarelo; as unidades diferentes são sombreadas em vermelho; as comuns a subgrupos de proteínas são sombreadas em azul ou verde. [De S. B. Carroll, J. K. Grenier, and S. D. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*, 2nd ed. Blackwell, 2005.]

verdade para outros genes instrumentais: uma fração significativa desses genes codifica fatores de transcrição que controlam a expressão de outros genes.

Muitos genes instrumentais (*toolkit*) codificam fatores de transcrição que regulam a expressão de outros genes.

Veremos como as proteínas Hox e outras proteínas instrumentais orquestram a expressão gênica um pouco mais adiante. Primeiro, há uma grande descoberta a descrever, revelando que o que aprendemos dos genes *Hox* tem implicações muito gerais para o reino animal.

Grupos de genes *Hox* controlam o desenvolvimento na maioria dos animais

Quando os genes homeobox foram descobertos em genes *Hox* de moscas, houve dúvida se essa característica tinha alguma peculiaridade com esses bizarros genes de mosca ou era mais amplamente distribuída, em outros insetos ou animais segmentados, por exemplo. Para abordar essa possibilidade, os pesquisadores procuraram homeoboxes nos genomas de outros insetos, bem como minhocas, sapos, vacas, e mesmo

humanos. Eles observaram muitos homeoboxes em cada um desses genomas animais.

As similaridades nas seqüências homeobox de espécies diferentes foram surpreendentes. Em mais de 60 dos aminoácidos do homeodomínio, algumas proteínas Hox de camundongos e sapos eram idênticas às seqüências das moscas em até 59 das 60 posições (Figura 12.9). Tendo em vista as grandes distâncias evolutivas entre esses animais, mais de 500 milhões de anos desde que tiveram um ancestral comum, a amplitude da similaridade das seqüências indica uma pressão muito forte para manter a seqüência do homeodomínio.

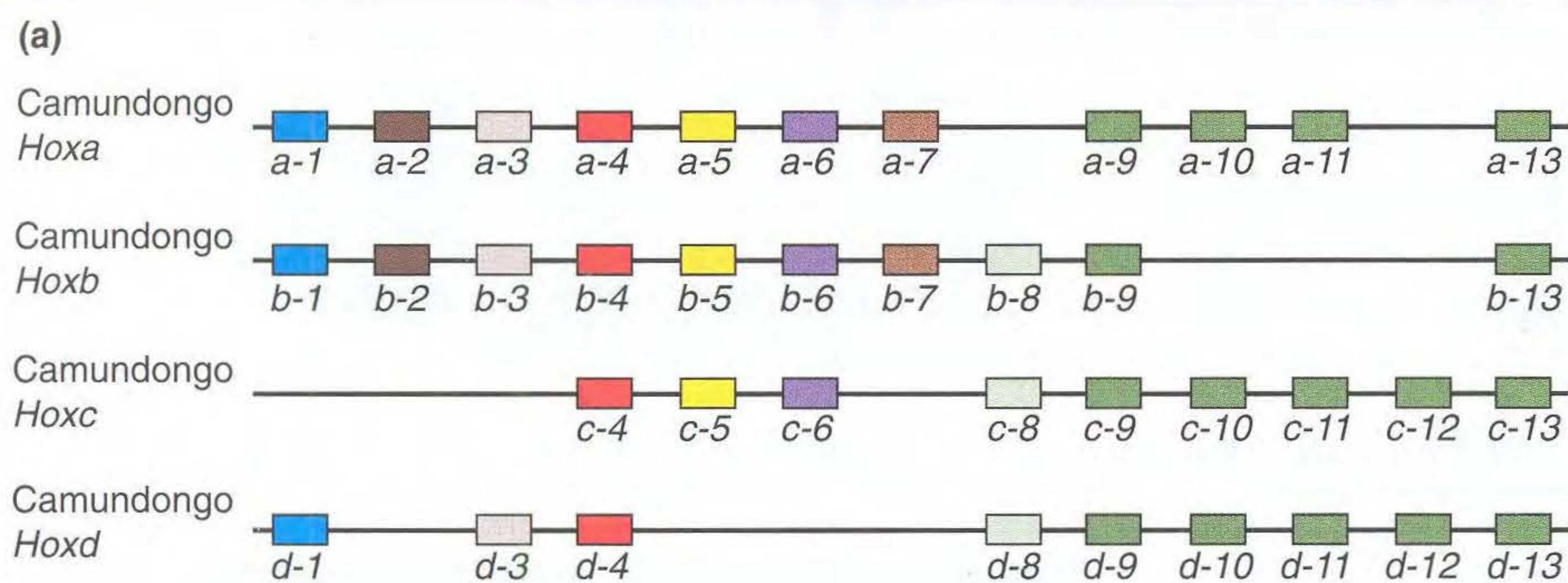
A existência dos genes *Hox* com homeoboxes ao longo do reino animal foi totalmente inesperada. Por que tipos diferentes de animais possuiriam os mesmos genes reguladores não era óbvio, motivo pelo qual os biólogos ficaram mais surpresos com os resultados quando foi examinada a organização e expressão de genes *Hox* em outros animais. Nos vertebrados, tais como o camundongo de laboratório, os genes *Hox* também estão aglomerados em quatro grandes complexos gênicos em quatro cromossomos diferentes. Cada grupo contém de 9 a 11 genes *Hox*, um total de 39 genes *Hox*. Além disso, a ordem dos genes nos complexos *Hox* de camundongos é paralela à ordem de suas contrapartes mais

A proteína Hox de *Drosophila* e a de vertebrados mostram marcantes similaridades

Mosca <i>Dfd</i>	PKRQRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKKDN	KLPNTKNVR
Anfíbio <i>Hox4</i>	TKRSRTAYTRQVLELEKEFHFNRYLTRRRRIEIAHSLGLTERQIKIWFQNRRMKWKKDN	RLPNTKTRS
Camundongo <i>HoxB4</i>	PKRSRTAYTRQVLELEKEFHFNRYLTRRRRIEIAHALCLSERQIKIWFQNRRMKWKKDH	KLPNTKIRS
Humano <i>HoxB4</i>	PKRSRTAYTRQVLELEKEFHFNRYLTRRRRIEIAHALCLSERQIKIWFQNRRMKWKKDH	KLPNTKIRS
Galinha <i>HoxB4</i>	PKRSRTAYTRQVLELEKEFHFNRYLTRRRRIEIAHSLCLSERQIKIWFQNRRMKWKKDH	KLPNTKIRS
Sapo <i>HoxB4</i>	AKRSRTAYTRQVLELEKEFHFNRYLTRRRRIEIAHTLRLSERQIKIWFQNRRMKWKKDH	KLPNTKIKS
Fugu <i>HoxB4</i>	PKRSRTAYTRQVLELEKEFHFNRYLTRRRRIEIAHTLCLSERQIKIWFQNRRMKWKKDH	KLPNTKIRS
Peixe-zebra <i>HoxB4</i>	AKRSRTAYTRQVLELEKEFHFNRYLTRRRRIEIAHTLRLSERQIKIWFQNRRMKWKKDH	KLPNTKIKS

FIG. 12.9 As seqüências do homeodomínio da proteína de *Drosophila* Deformed e de vários membros dos genes da proteína Hox grupo 4 são muito similares. As unidades em comum são sombreadas em amarelo. As unidades divergentes são sombreadas em vermelho; as unidades comuns a subgrupos de proteínas são sombreadas em azul. As regiões flangeadoras C-terminais muito similares fora do homeodomínio são sombreadas em verde. [De S. B. Carroll, J. K. Grenier, and S. D. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*, 2nd ed. Blackwell, 2005.]

A ordem dos genes *Hox* é paralela à ordem das partes do corpo nas quais eles são expressos



(b)
Embrião de camundongo

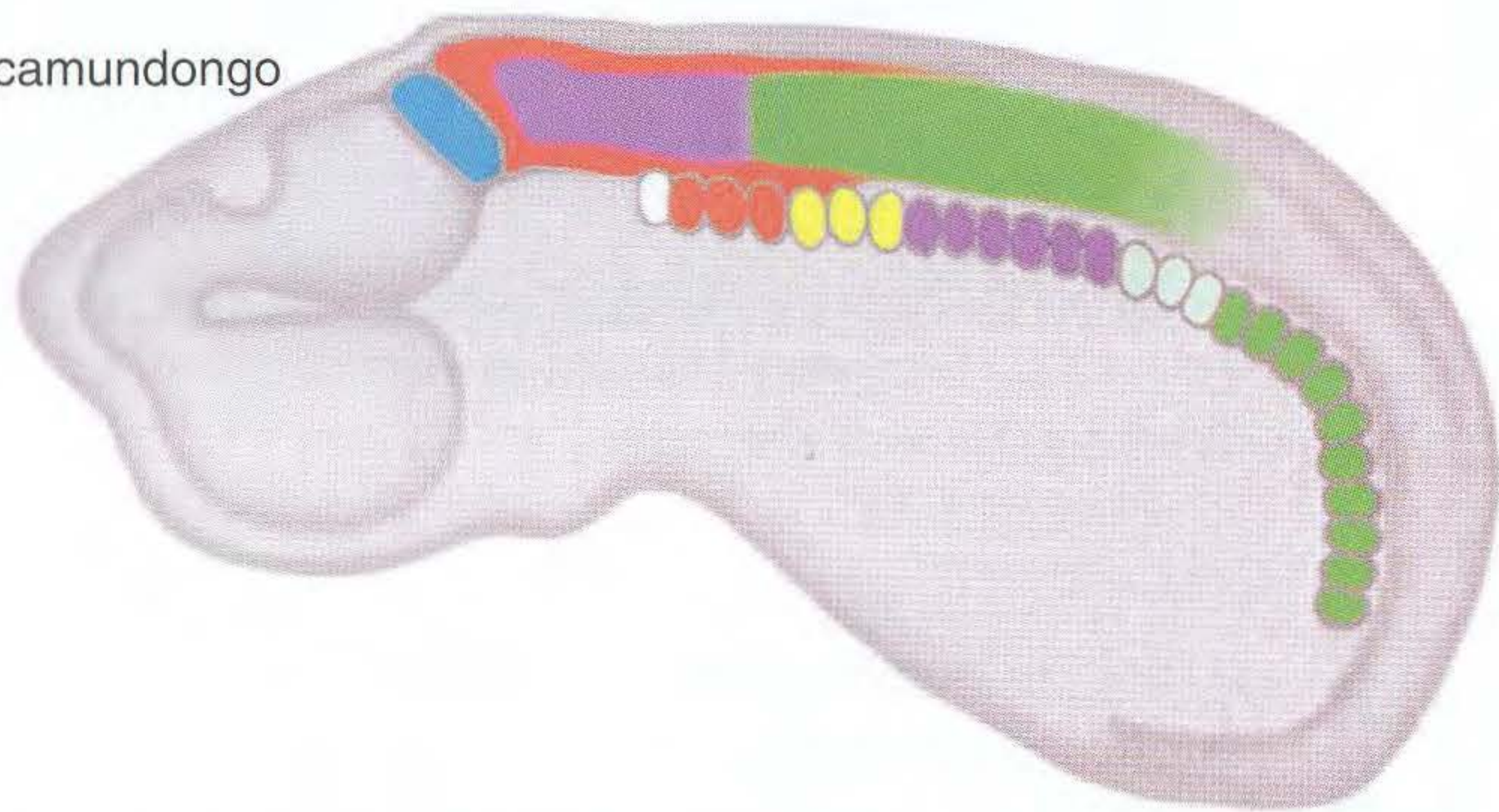


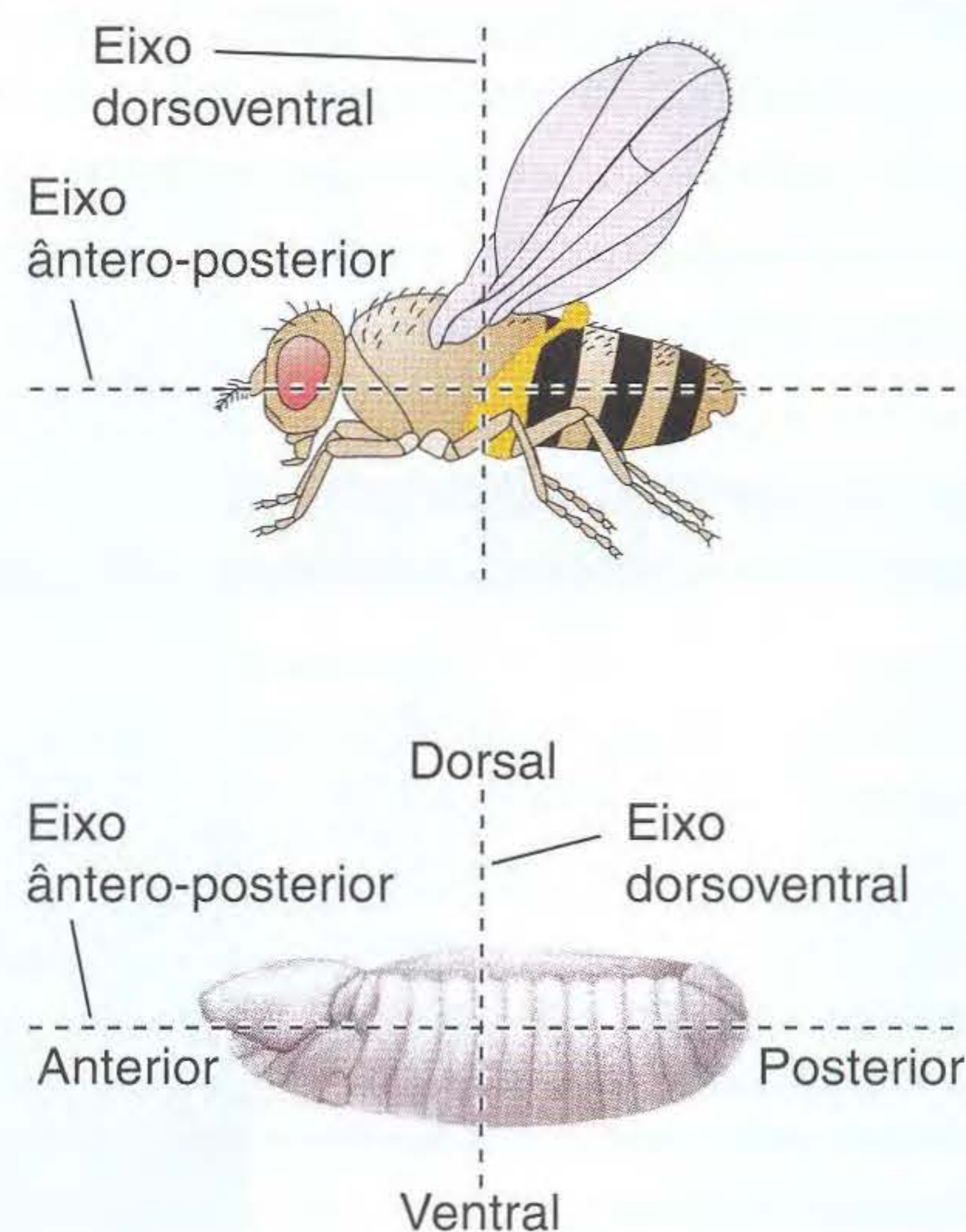
FIG. 12.10 Como o das moscas-das-frutas, os genes *Hox* de vertebrados são organizados em grupos e expressos ao longo do eixo ântero-posterior. (a) No camundongo, quatro complexos dos genes *Hox*, compreendendo 39 genes no total, estão presentes em quatro cromossomos diferentes. Nem todo gene é representado em cada complexo. Alguns foram perdidos no curso da evolução. (b) Os genes *Hox* são expressos em domínios distintos ao longo do eixo ântero-posterior do embrião de camundongo. As cores representam os grupos diferentes de genes mostrados na parte a. [De S. B. Carroll, "Homeotic Genes and the Evolution of Arthropods and Chordates," *Nature* 376, 1995, 479-485.]

relacionadas nos complexos *Hox* das moscas, bem como em cada um dos outros grupos *Hox* de camundongos (Figura 12.10a). Essa correspondência indica que os complexos *Hox* de insetos e vertebrados estão relacionados, e que alguma forma de complexo *Hox* existiu em seu distante ancestral comum. Os quatro complexos *Hox* em camundongos surgiram por duplicações de complexos *Hox* inteiros (talvez de cromossomos inteiros) nos ancestrais dos vertebrados.

Por que tais animais diferentes têm esses conjuntos de genes em comum? Seus ancestrais comuns indicam que os genes *Hox* têm algum papel fundamental no desenvolvimento da maioria dos animais. Esse papel é aparente pelas análises de como os genes *Hox* são expressos em animais diferentes. Nos embriões dos vertebrados, genes *Hox* adjacentes também são expressos em domínios adjacentes ou parcialmente superpostos ao longo do eixo ântero-posterior. Além disso, a ordem dos genes *Hox* nos complexos corresponde à ordem da cabeça a cauda das regiões do corpo nas quais os genes são expressos (Figura 12.10b).

Os padrões de expressão dos genes *Hox* dos vertebrados sugeriram que eles também especificam a identidade das regiões corpóreas, e subseqüentes análises de mutantes do gene *Hox* levantaram essa sugestão. Por exemplo, as mutações nos genes *Hoxa11* e *Hoxd11* causam a transformação homeótica das vértebras sacrais em vértebras lombares (Figura 12.11). Assim, como na mosca, a perda ou ganho de função dos genes *Hox* em vertebrados causa a transformação da identidade de

estruturas repetidas em série. Tais resultados foram obtidos em várias classes, incluindo mamíferos, aves, anfíbios e peixes. Além disso, os grupos de genes *Hox* foram mostrados governando a padronização de outros insetos e dispostos em regiões ao longo do eixo ântero-posterior em anelídeos, moluscos, nematódeos, vários artrópodes, cordados primitivos platelmintos e outros animais. Portanto, a despeito das



As correlações entre os eixos corpóreos embrionário e adulto

Os genes *Hox* regulam a identidade de estruturas serialmente repetidas em vertebrados

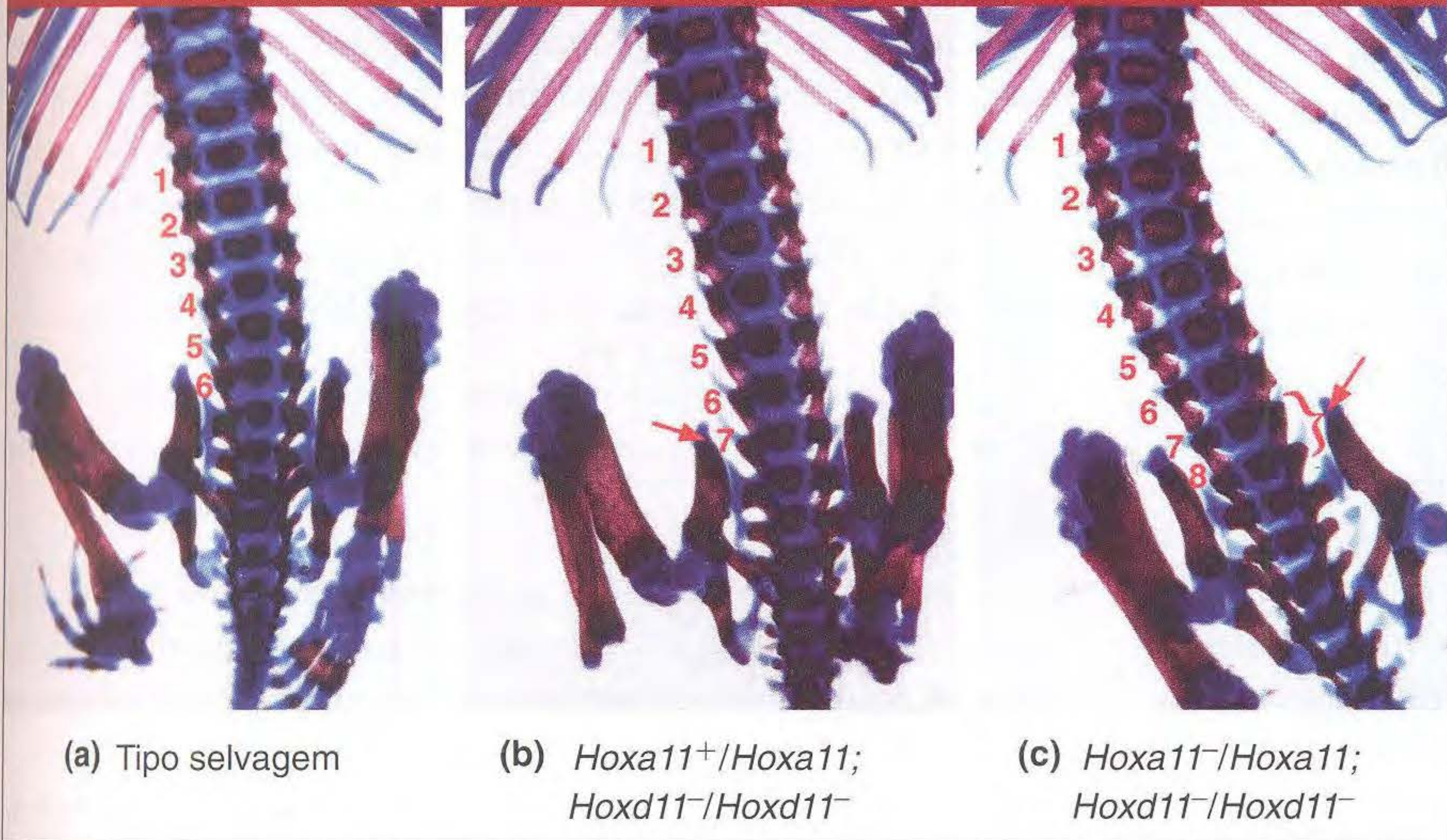


FIG. 12.11 As morfologias das diferentes regiões da coluna vertebral são reguladas pelos genes *Hox*. (a) No camundongo, seis vértebras lombares formam o anterior da vértebra sacral (números em vermelho). (b) Em camundongos sem a função de posterioridade do gene *Hoxd11* e possuindo uma cópia funcional do gene *Hoxa11*, formam-se sete vértebras lombares e uma vértebra sacral é perdida. (c) Em camundongos sem a função de ambos *Hoxa11* e *Hoxd11*, formam-se oito vértebras lombares e duas vértebras sacrais são perdidas. [Fotos por cortesia da Dra. Anne Boulet, HHMI, University of Utah; de S. B. Carroll, J. K. Grenier, and S. D. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*, 2nd ed. Blackwell, 2005.]

enormes diferenças na anatomia, a posse de um ou mais grupos de genes *Hox* que estão dispostos em regiões ao longo do eixo corpóreo principal é uma característica comum, fundamental, de pelo menos todos os animais bilaterais. De fato, as lições surpreendentes dos genes *Hox* mostraram o que se tornou uma tendência geral entre os genes instrumentais; isto é, a maioria dos genes instrumentais são comuns a animais diferentes.

A despeito de grandes diferenças em anatomia, muitos genes instrumentais são comuns a uma ampla gama de diferentes ramos de animais.

Façamos agora um inventário do resto do instrumental para ver que outros princípios gerais emergem.

12.3 Definição de Todas as Ferramentas

Os genes *Hox* são talvez os membros mais bem conhecidos das ferramentas, mas são uma pequena família em um grupo muito grande de genes necessários para o desenvolvimento dos números apropriados, formas, tamanhos e tipos de parte corpóreas. Pouco se sabia sobre o resto das ferramentas até finais de 1970 e início dos anos de 1980, quando Christine Nüsslein-Volhard e Eric Wieschaus, que trabalhavam no Max Planck Institute em Tübingen, Alemanha, encontraram os genes necessários para a formação da organização segmentar do embrião e larva de *Drosophila*.

Até seus esforços, a maioria dos trabalhos com o desenvolvimento da mosca enfocavam fenótipos adultos viáveis e não embriões. Nüsslein-Volhard e Wieschaus perceberam que os tipos de genes que eles estavam procurando eram provavelmente letais para os embriões ou larvas em mutantes homozigotos. Assim, eles criaram um esquema para procurar genes que são necessários no **zigoto** (o produto da fertilização; Figura 12.12, embaixo). Eles também desenvolveram tria-

gens para identificar os genes com produtos que funcionam no zigoto, antes que o genoma zigótico seja ativado, e que são necessários para a padronização do embrião. Os genes com produtos dados pela fêmea ao zigoto são chamados de **genes de efeito materno**. Os fenótipos mutantes de genes de estrito efeito materno dependem apenas do genótipo da fêmea (Figura 12.12, em cima).

Nessas triagens, foram identificados genes que eram necessários para fazer o número e padrões apropriados dos segmentos larvares, para fazer suas três camadas de tecidos e padronizar os detalhes finos da anatomia animal. O poder das triagens genéticas era sua natureza sistemática. Saturando cada cromossomo da mosca (exceto o pequeno quarto cromossomo) com mutações induzidas quimicamente, os pesquisa-

Triagens genéticas para genes ferramentais necessários materna e zigoticamente

GENES NECESSÁRIOS MATERNAMENTE

Genitores	Prole
$m/+ \text{ ♂} \times m/+ \text{ ♀}$	$m/m, m/+, +/+$ todos normais
$m/m \text{ ♂} \times m/+ \text{ ♀}$	$m/m, m/+$ todos normais
$+/+, m/+, \text{ or } m/m \text{ ♂} \times m/m \text{ ♀}$	$m/+, m/m$ todos com fenótipo mutante

GENES NECESSÁRIOS ZIGOTICAMENTE

Genitores	Prole
$m/+ \text{ ♂} \times m/+ \text{ ♀}$	$m/+, +/+$ normal
	m/m fenótipo mutante

FIG. 12.12 As triagens genéticas identificam se um produto gênico funciona no ovócito ou no zigoto. Os fenótipos da prole dependem do genótipo materno (*em cima*) para genes de efeito materno ou (*embaixo*) do genótipo da prole (zigótico) para genes zigoticamente necessários (*m*, mutante; *+*, tipo selvagem).

dores foram capazes de identificar a maioria dos genes que eram necessários para a construção da mosca. Por seus esforços pioneiros, Nüsslein-Volhard, Wieschaus e Lewis dividiram o Prêmio Nobel de 1995 de Medicina ou Fisiologia.

A mais marcante das características dos mutantes recém-identificados foi que eles apresentavam defeitos marcantes, mas discretos, na organização ou padronização do embrião. Isto é, a larva morta não era uma carcaça amorfa, mas exibia defeitos de padronização específicos, e geralmente marcantes. O corpo da larva de *Drosophila* tem várias características cujos números, posição ou padrão podem servir como marcos para diagnosticar ou classificar as anomalias nos animais mutantes. Cada locus podia, portanto, ser classificado de acordo com o eixo corpóreo que é afetado e o padrão de defeitos causados pelas mutações. Os cruzamentos revelam se o locus é ativo no ovócito materno ou zigoto. Cada classe de genes parece representar etapas diferentes no refinamento progressivo do plano corpóreo embrionário, dos que afetam grandes regiões do embrião até aqueles com áreas mais limitadas de influência.

Para qualquer gene ferramental, três partes da informação são fundamentais para compreender o funcionamento gênico: (1) o fenótipo mutante; (2) o padrão da expressão gênica; e (3) a natureza do produto gênico. O estudo amplo de algumas dezenas de genes levou a um quadro detalhado de como cada eixo do corpo é estabelecido e subdividido em segmentos de camadas germinativas.

Os eixos ântero-posterior e dorsoventral

Algumas dúzias de genes são necessárias para a organização apropriada do eixo corpóreo ântero-posterior do embrião da

mosca. Os genes são agrupados em cinco classes distintas com base em sua área de influência no padrão embrionário.

- A primeira classe estabelece o eixo ântero-posterior e consiste nos genes de efeito materno. Um membro importante dessa classe é o gene *Bicoid*. Os embriões das mães dos mutantes *Bicoid* não apresentam a região anterior do embrião (Figura 12.13), significando isso que o gene é necessário para o desenvolvimento dessa região.

As três classes seguintes são genes zigoticamente ativos necessários para o desenvolvimento de segmentos do embrião.

- A segunda classe contém os **genes gap**. Cada um desses genes afeta a formação de um bloco contíguo de segmentos; as mutações nos genes gap levam a grandes espaços na segmentação (Figura 12.14).
- A terceira classe constitui os **genes de regra dos pares**, que atuam com uma periodicidade de segmento duplo. Os mutantes de regras dos pares são partes que faltam de cada par de segmentos, mas genes diferentes de regra dos pares afetam partes diferentes de cada segmento duplo. Por exemplo, o gene *even-skipped* afeta um grupo de limites segmentares, e o gene *odd-skipped* afeta o conjunto complementar de limites (Figura 12.14).
- A quarta classe consiste nos **genes de polaridade segmentar**, que afeta a padronização dentro de cada segmento. Os mutantes dessa classe apresentam defeitos na polaridade e número de segmentos (Figura 12.14).

A quinta classe de genes determina o destino de cada segmento.

- A quinta classe inclui os genes *Hox* já discutidos; os mutantes *Hox* não afetam o número de segmentos, mas alteram o aspecto de um ou mais segmentos.

O eixo dorsoventral também é subdividido em regiões. Vários genes materno-fetais, tais como o *dorsal*, são necessários para iniciar o estabelecimento da polaridade dorsoventral normal do embrião. Os mutantes *Dorsal* são “dorsalizados” e não têm estruturas ventrais (tais como o mesoderma e sistema nervoso). Muitos genes zigoticamente ativos também são necessários para a subdivisão do eixo dorsoventral.

Expressão dos genes ferramentais

Para compreender a relação entre genes e o fenótipo mutante, devemos conhecer a época e localização dos padrões de expressão gênica e a natureza molecular dos produtos gênicos. Os padrões de expressão dos genes ferramentais correspondem vivamente a seus fenótipos, pois são precisamente correlacionados com as partes do corpo em desenvolvimento que são alteradas nos mutantes. Cada gene é expresso em uma região que pode ser mapeada em coordenadas específicas ao longo do eixo do embrião. Por exemplo, o efeito materno da proteína Bicoid é expresso em um padrão graduado que emana do pólo anterior do embrião inicial, a seção do embrião que falta nos mutantes (Figura 12.15a).

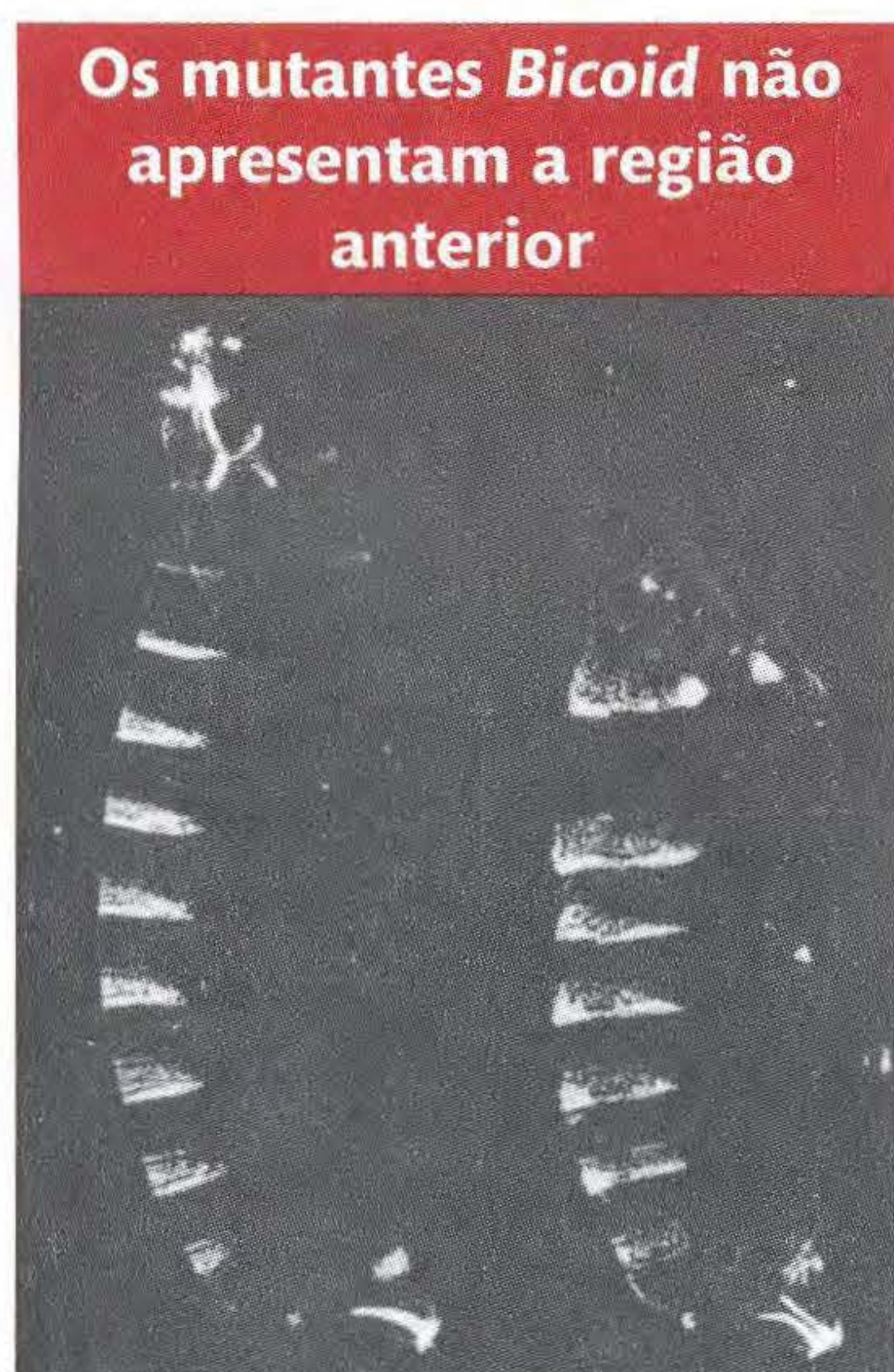


FIG. 12.13 O gene de efeito materno *Bicoid* (*bcd*) afeta a parte anterior da larva em desenvolvimento. Estas fotomicrografias são de larvas de *Drosophila* que foram preparadas para mostrar seus rígidos exoesqueletos. Estruturas densas, tais como as bandas de dentículos segmentais, aparecem em branco. (Esquerda) Uma larva normal. (Direita) Uma larva de um mutante homozigoto *bcd* fêmea. As estruturas da cabeça e torácica anterior estão faltando. [De C. H. Nüsslein-Volhard, G. Frohnhofer, and R. Lehmann, "Determination of Anteroposterior Polarity in *Drosophila*," *Science* 238, 1987, 1678.]

Os mutantes de genes de segmentação têm falta de partes de segmentos

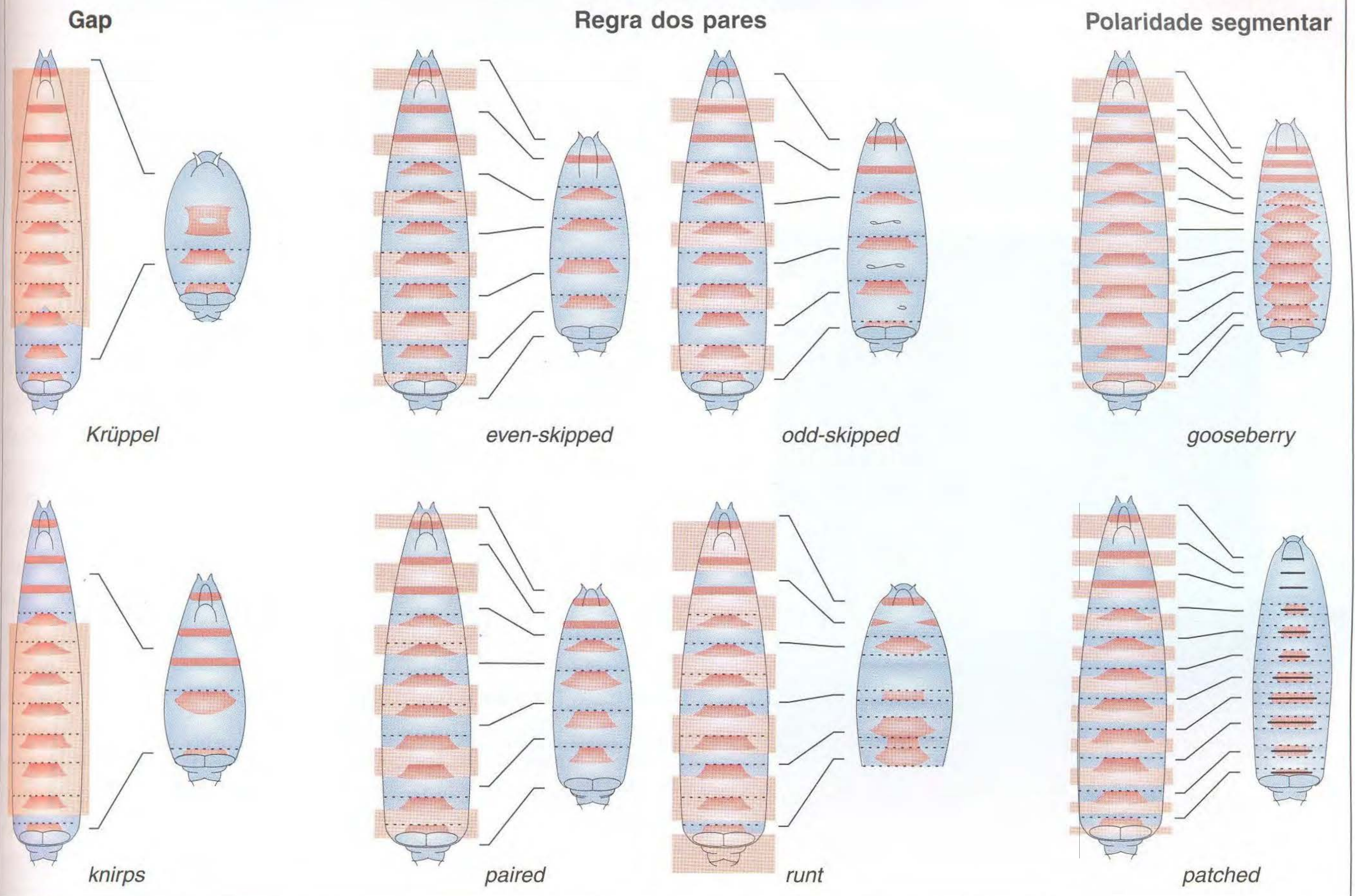


FIG. 12.14 Classes de mutantes de gene de segmentação de *Drosophila*. Estes diagramas mostram mutantes gap, regra dos pares e polaridade segmentar. Os trapézios em vermelho são as bandas densas de exoesqueleto vistas na Figura 12.13. O limite de cada segmento é indicado por uma linha pontilhada. O diagrama à esquerda de cada par mostra uma larva tipo selvagem, e o diagrama à direita mostra o padrão formado em determinado mutante. As regiões sombreadas em rosa nos diagramas tipo selvagem indicam os domínios da larva que estão faltando ou estão afetados no mutante.

Similarmente, as proteínas gap são expressas em blocos de células que correspondem a futuras posições dos segmentos que estão faltando nos respectivos genes gap (Figura 12.15b). As proteínas de regra dos pares são expressas em padrões de faixas: uma faixa transversa é expressa a cada 2 segmentos, em um total de 7 faixas cobrindo os 14 futuros segmentos corpóreos (a posição e periodicidade das faixas correspondem à periodicidade de defeitos nas larvas mutantes), como mostrado na Figura 12.15c. Muitos genes de polaridade segmentar são expressos em faixas de células dentro de cada segmento, 14 faixas no total (Figura 12.15d). Note que os domínios da expressão gênica tornam-se progressivamente mais refinados à medida que o desenvolvimento continua: os genes são expressos primeiro em grandes regiões (proteínas gap), então em faixas de três a quatro células de largura (proteínas de regra dos pares) e, a seguir, em faixas de uma a duas células de largura (proteínas de polaridade segmentar).

Entre os genes de padrão do eixo dorsoventral, a proteína Dorsal é expressa em um gradiente ao longo do eixo dorsoventral, que é o mais alto nível de acúmulo nas células ven-

trais (Figura 12.16a). Vários genes zigóticos de padrão dorsoventral são expressos em sub-regiões diferentes ao longo do eixo dorsoventral, correspondendo a regiões que dão origem a camadas tissulares particulares, tais como o mesoderma e o neuroectoderma (a parte do ectoderma que dá origem ao sistema nervoso ventral), como mostrado na Figura 12.16b.

Em adição ao que aprendemos dos padrões espaciais da expressão dos genes ferramentais, a ordem de expressão desses genes ao longo do tempo é lógica. O gene materno Bicoid expresso aparece antes das proteínas gap zigóticas, que são expressas antes dos padrões de 7 faixas das proteínas de regra dos pares aparecer, que, por sua vez, precede os padrões de 14 faixas das proteínas de polaridade segmentar. Similarmente, o gradiente da proteína Dorsal forma-se antes das faixas longitudinais dos genes de padronagem do eixo dorsoventral zigótico. A ordem de expressão do refinamento progressivo dos domínios dentro do embrião revela que a produção do plano corpóreo é um processo gradativo, com importantes subdivisões do corpo destacadas primeiro e, então, refinadas até ser estabelecido um padrão fino. A ordem de ação gênica

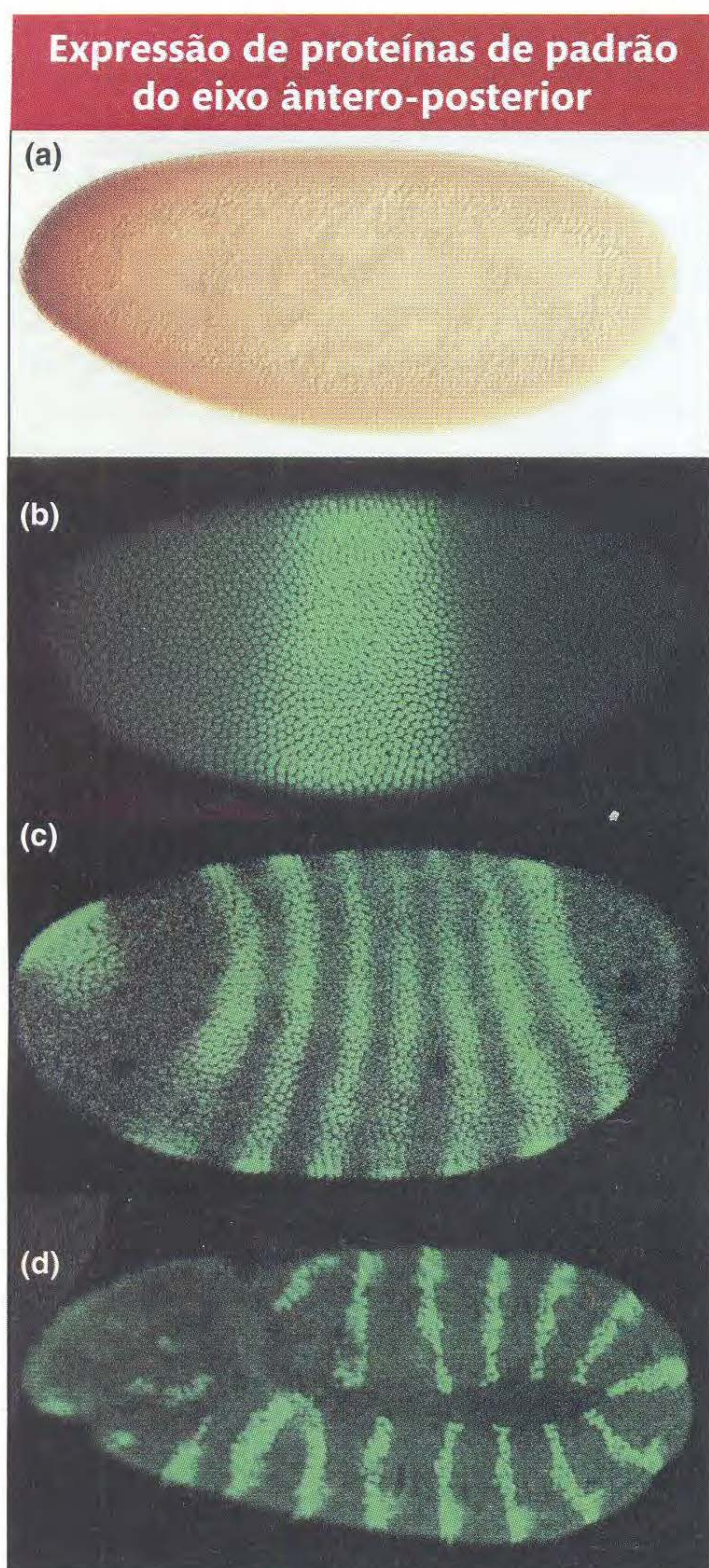


FIG. 12.15 Os padrões de expressão de genes ferramentais correspondem a fenótipos mutantes. Os embriões de *Drosophila* foram corados com anticorpos para (a) a proteína Bicoid maternamente derivada, (b) proteína gap Krüppel, (c) proteína Hairy de regra dos pares e (d) proteína de polaridade segmentar Engrailed e visualizada por métodos imunoenzimáticos (a) ou de imunofluorescência (b-d). Cada proteína está localizada nos núcleos em regiões do embrião que são afetadas por mutações nos respectivos genes. [Fotomicrografias por cortesia de (a) Ruth Lehmann e (b-d) James Langeland.]

sugere, ainda, que a expressão de um grupo de genes pode controlar a expressão dos grupos de genes seguintes.

Um indício de que essa progressão é de fato o caso vem da análise dos efeitos das mutações nos genes ferramentais na expressão de outros genes ferramentais. Por exemplo, nos embriões de mães mutantes *Bicoid*, a expressão de vários genes gap é alterada, bem como os genes de regra dos pares e de polaridade segmentar. Esse achado sugere que a proteína Bicoid de algum modo (direta ou indiretamente) influencia a regulação dos genes gap.

Um outro indício de que a expressão de um grupo de genes pode controlar a expressão de sucessivos grupos de genes vem do exame dos produtos proteicos. A inspeção da sequência proteica Bicoid revela que ela contém um homeodomínio, correlato mas distinto dos das proteínas Hox. Assim, Bicoid tem as propriedades de um fator de transcrição de ligação ao DNA. Cada gene gap também codifica um fator de transcrição, como cada gene de regra dos pares, vários genes de polaridade segmentar, o gene *dorsal*, e vários genes de padronização do eixo dorsoventral. Esses fatores de transcrição incluem representantes da maioria das famílias conhecidas de proteínas de ligação ao DNA específicas de sequências. Assim, não há restrição quanto a que famílias eles podem pertencer, muitas proteínas ferramentais de ação precoce são fatores de transcrição. As que não são fatores de transcrição tendem a ser componentes de vias de sinalização (Quadro 12.1). Essas vias, mostradas em forma genérica na Figura 12.17, medeiam processos de sinalização induzidos por ligandos entre células, e seu produto geralmente leva à ativação gênica ou repressão. Assim, a maioria das proteínas ferramentais, seja direta (como fatores de transcrição) ou indiretamente (como componentes de vias de sinalização), afetam a regulação gênica.

Muitas proteínas ferramentais são fatores de transcrição ou componentes de vias de transdução de sinal mediadas por ligando.

O controle genético do desenvolvimento é, então, fundamentalmente uma questão de regulação gênica no espaço e no tempo. Como ligar e desligar genes ferramentais cons-

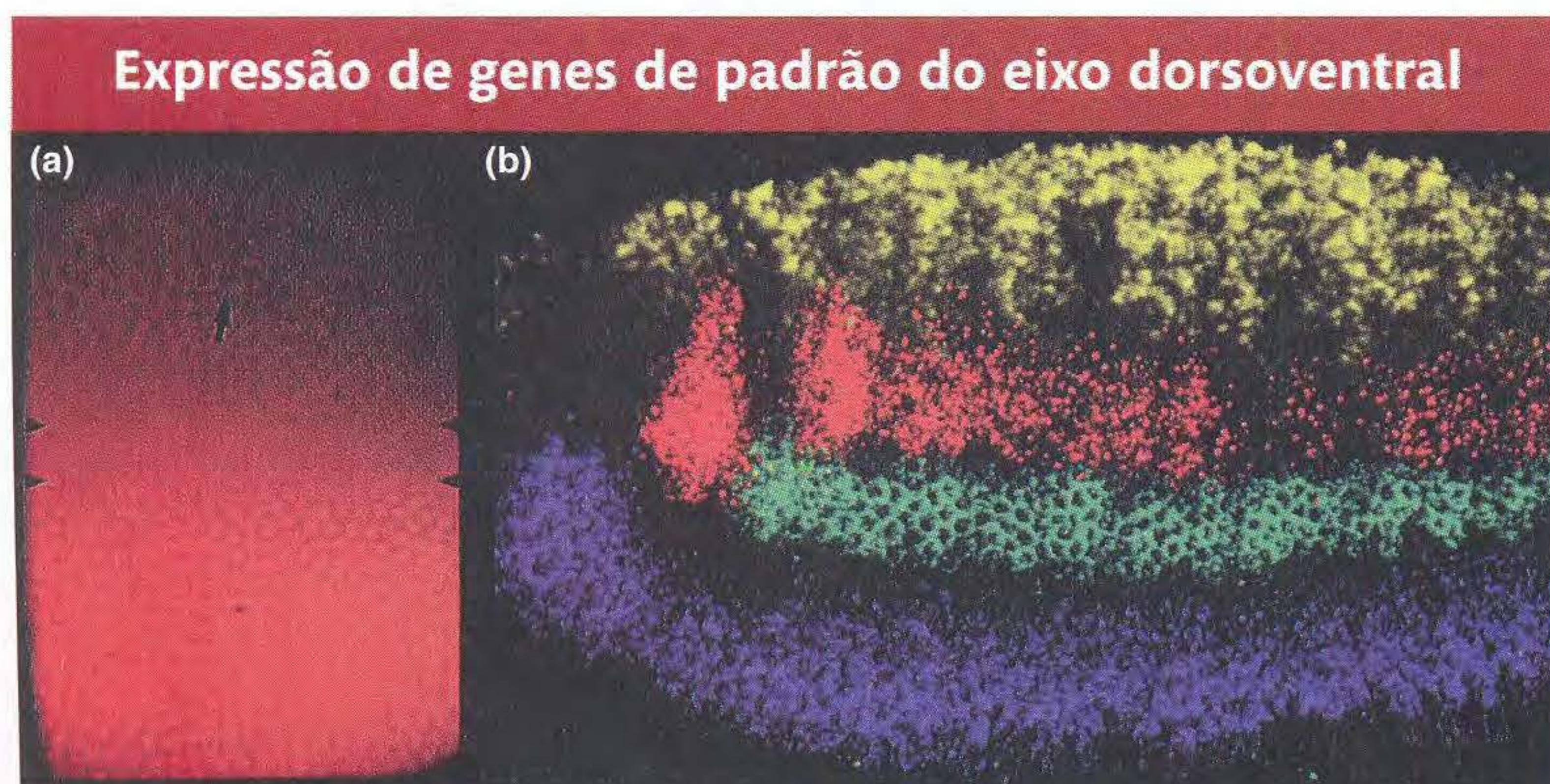


FIG. 12.16 A expressão dos genes de padrão do eixo dorsoventral corresponde a camadas particulares de tecidos. (a) A proteína Dorsal derivada maternamente é expressa em um gradiente, com a maior concentração de Dorsal nos núcleos de células ventrais (embaixo, na foto). (b) A expressão dos quatro genes zigóticos de padrão do eixo dorsoventral revelada pela hibridização *in situ* ao RNA. Nessa visão lateral, são revelados os domínios dos genes *decapentaplegic* (amarelo), *muscle segment homeobox* (vermelho), *intermediate neuroblast defective* (verde), e *ventral neuroblast defective* (azul). [Fotos por cortesia de (a) Michael Levine e (b) David Kosman, Bill McGinnis, and Ethan Bier.]

Quadro 12.1 Exemplos de Genes do Eixo A-P de *Drosophila* que Contribuem para a Formação de Padrão

Símbolo gênico	Nome do gene	Função da proteína	Papel(éis) no desenvolvimento inicial
<i>hb-z</i>	<i>hunchback-zygotic</i>	Fator de transcrição – proteína zinc-finger	Gene gap
<i>Kr</i>	<i>Krüppel</i>	Fator de transcrição – proteína zinc-finger	Gene gap
<i>kni</i>	<i>knirps</i>	Fator de transcrição – proteína tipo receptor de esteróide	Gene gap
<i>eve</i>	<i>even-skipped</i>	Fator de transcrição – proteína homeodomínio	Gene de regra dos pares
<i>ftz</i>	<i>fushi tarazu</i>	Fator de transcrição – proteína homeodomínio	Gene de regra dos pares
<i>opa</i>	<i>odd-paired</i>	Fator de transcrição – proteína zinc-finger	Gene de regra dos pares
<i>prd</i>	<i>paired</i>	Fator de transcrição – proteína PHOX	Gene de regra dos pares
<i>en</i>	<i>engrailed</i>	Fator de transcrição – proteína homeodomínio	Gene de polaridade segmentar
<i>ci</i>	<i>cubitus-interruptus</i>	Fator de transcrição – proteína zinc-finger	Gene de polaridade segmentar
<i>wg</i>	<i>wingless</i>	Sinalização da proteína WG	Gene de polaridade segmentar
<i>hh</i>	<i>hedgehog</i>	Proteína de sinalização HH	Gene de polaridade segmentar
<i>fu</i>	<i>fused</i>	Cinase de serina/treonina citoplasmática	Gene de polaridade segmentar
<i>ptc</i>	<i>patched</i>	Proteína transmembranar	Gene de polaridade segmentar
<i>arm</i>	<i>armadillo</i>	Proteína de junção célula-célula	Gene de polaridade segmentar
<i>lab</i>	<i>labial</i>	Fator de transcrição – proteína homeodomínio	Gene de identidade segmentar
<i>Dfd</i>	<i>Deformed</i>	Fator de transcrição – proteína homeodomínio	Gene de identidade segmentar
<i>Antp</i>	<i>Antennapedia</i>	Fator de transcrição – proteína homeodomínio	Gene de identidade segmentar
<i>Ubx</i>	<i>Ultrabithorax</i>	Fator de transcrição – proteína homeodomínio	Gene de identidade segmentar

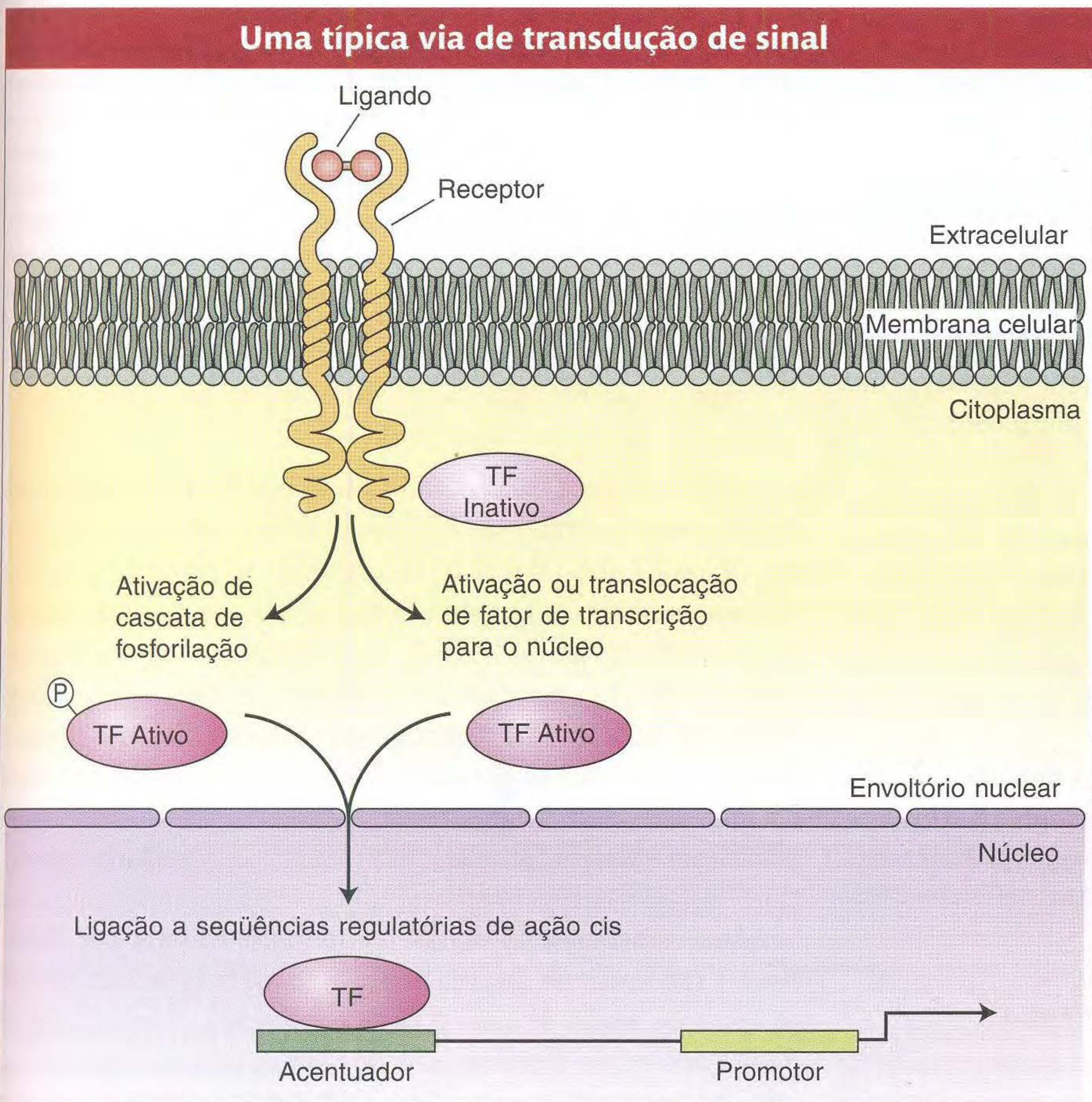


FIG. 12.17 A maioria das vias de sinalização operam por uma lógica similar, mas têm componentes proteicos diferentes e mecanismos de transdução de sinal. A sinalização começa quando um ligando se une a um receptor ligado a membrana, levando à liberação ou ativação de proteínas intracelulares. A ativação de receptor geralmente se liga à modificação de fatores de transcrição inativos (TF). Os fatores de transcrição modificados são translocados para o núcleo celular, onde eles se ligam a seqüências de DNA regulatórias de ação cis ou a proteínas de ligação ao DNA e regulam o nível de transcrição do gene-alvo. [De S.B. Carroll, J. K. Grenier, and S. D. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. 2nd ed. Blackwell, 2005.]

troem a forma animal? E como isso é coreografado durante o desenvolvimento? Para responder a essas perguntas, examinaremos as interações entre as proteínas ferramentais da mosca e os genes em maiores detalhes. Os mecanismos que veremos para o controle da expressão de genes ferramentais no embrião de *Drosophila* emergiram como modelos para a regulação espacial da expressão gênica no desenvolvimento animal em geral.

► 12.4 Regulação Espacial da Expressão Gênica no Desenvolvimento

Vimos que os genes ferramentais são expressos em relação a coordenadas no embrião. Mas como as coordenadas espaciais do embrião em desenvolvimento funcionam como instruções para os genes, para ligá-los ou desligá-los em padrões precisos? Como descrito nos Capítulos 10 e 11, o controle fisiológico da expressão gênica em bactérias e eucariontes simples é finalmente governado por proteínas de ligação a seqüências específicas de ligação ao DNA agindo em elementos regulatórios de ação cis (por exemplo, operadores e elementos de seqüência de ativação antecedente, ou UAS). Similarmente, o controle espacial da expressão gênica durante o desenvolvimento é amplamente governado pela interação de fatores de transcrição com elementos regulatórios de ação cis. Entretanto, o controle espacial e temporal da regulação gênica no desenvolvimento de um embrião multicelular tridimensional requer a ação de mais fatores de transcrição em elementos mais numerosos e mais complexos de ação cis.

Para definir uma posição em um embrião, deve existir uma informação regulatória capaz de distinguir essa posição das regiões adjacentes. Se representarmos um embrião tridimensional como um globo, então a **informação posicional** deve ser especificada indicando a longitude (local ao longo do eixo ântero-posterior), latitude (local ao longo do eixo dorsoventral) e altitude ou profundidade (posição nas camadas germinativas). Ilustraremos os princípios gerais de como as posições da expressão gênica são especificadas com três exemplos. Esses exemplos devem ser pensados apenas como algumas ilustrações do grande número de interações regulatórias que governam o desenvolvimento da mosca e do animal. O desenvolvimento é um contínuo no qual cada padrão de atividade gênica tem uma base causal precedente. Todo o processo inclui dezenas de milhares de interações regulatórias e produtos.

Enfocaremos algumas conexões entre os genes em níveis diferentes de hierarquias que determinam o plano corpóreo segmental básico, e os *pontos nodais* onde genes importantes integram vários estímulos regulatórios e respondem produzindo expressões gênicas mais simples.

Gradientes maternos e ativação gênica

A proteína Bicoid é um fator de transcrição tipo homeodomínio que é traduzido do mRNA depositado no zigoto e situado

Os genes gap são ativados por proteínas específicas fornecidas maternamente

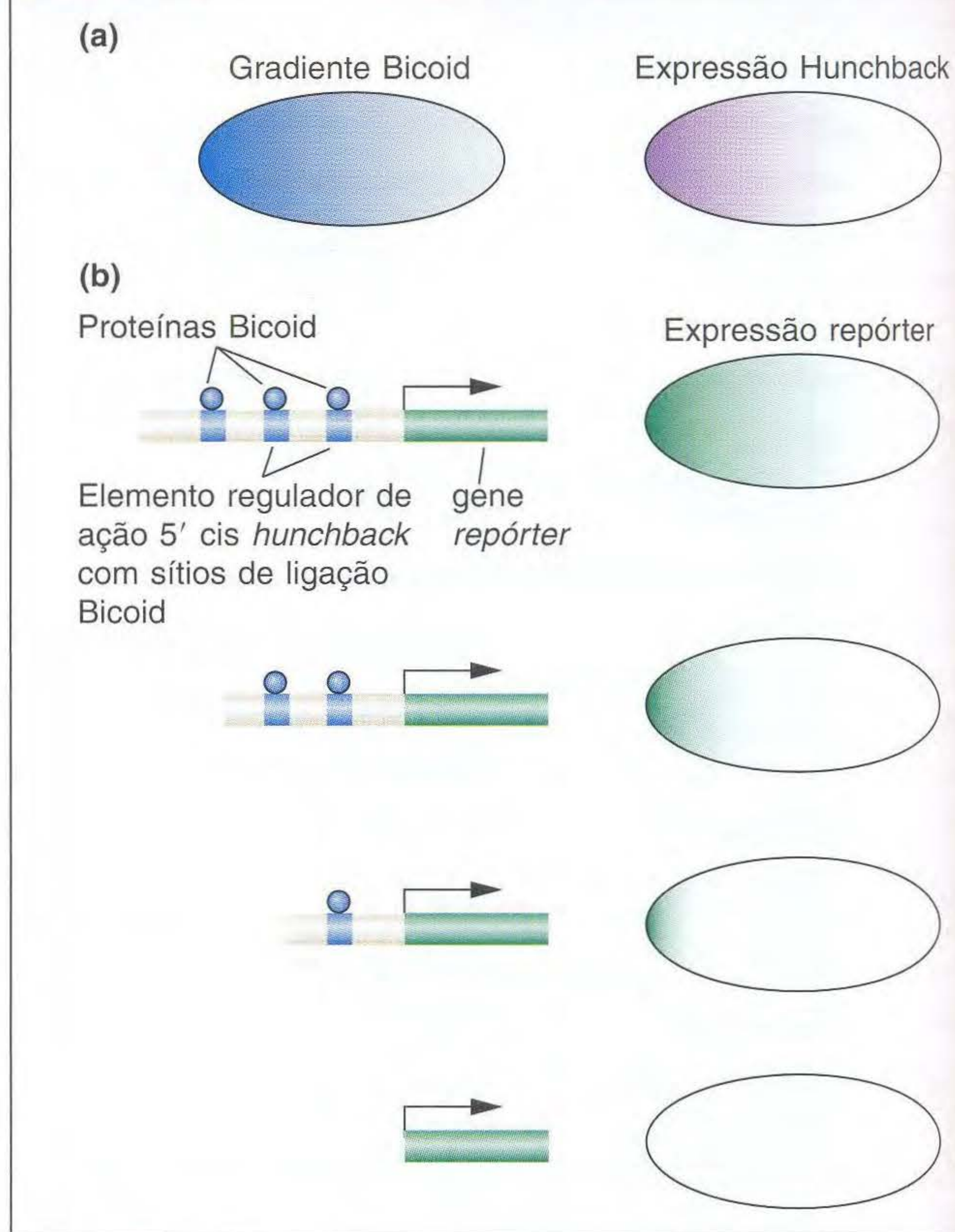


FIG. 12.18 A proteína Bicoid ativa a expressão zigótica do gene *hunchback*. (a) A expressão da proteína Bicoid é graduada ao longo do eixo ântero-posterior. O gene gap *hunchback* é expresso na metade anterior do zigoto. (b) A proteína Bicoid (azul) liga-se a três sítios 5' do gene *hunchback*. Quando esse DNA 5' é colocado antecedendo um gene repórter, a expressão do gene repórter recapitula o padrão da expressão *hunchback* (em cima, à direita). Entretanto, a deleção progressiva de um, dois ou de todos os três sítios Bicoid ou leva a uma expressão mais restrita do gene repórter, ou o suprime. Essas observações mostram que o nível e padrão da expressão *hunchback* são controlados por Bicoid por sua ligação a seqüências reguladoras de DNA *hunchback*.

no pólo anterior. Como o embrião inicial de *Drosophila* é um sincício, sem membranas celulares que possam impedir a difusão de moléculas proteicas, Bicoid pode se difundir pelo citoplasma. Essa difusão estabelece um gradiente de concentração de proteínas (Figura 12.18a): a proteína Bicoid é altamente concentrada na extremidade anterior, e essa concentração diminui gradualmente à medida que a distância dessa extremidade aumenta, até haver muito pouca proteína Bicoid no meio do embrião. Esse gradiente de concentração dá informação posicional sobre a localização ao longo do eixo ântero-posterior. Uma alta concentração significa a extremidade anterior, e a menor concentração significa a metade, e assim por diante. Assim, um modo de garantir que um gene seja ativado em um só local ao longo do eixo é ligar a expressão gênica ao nível de concentração. Seria o caso dos genes gap, que devem ser ativados em regiões específicas ao longo do eixo.

Vários genes zigóticos, incluindo os genes gap, são regulados por níveis diferentes da proteína Bicoid. Por exemplo, o gene *hunchback* é um gene gap ativado no zigoto na metade anterior do embrião. Essa ativação é pela ligação direta da proteína Bicoid a três sítios 5' do promotor do gene *hunchback*. Bicoid liga-se aos sítios *cooperativamente*; isto é, a ligação da molécula de proteína Bicoid a um sítio facilita a ligação de outras moléculas Bicoid a sítios vizinhos.

Como a ativação de *hunchback* depende do gradiente de concentração pode ser visto fazendo alguns testes *in vivo*. Esses testes requerem a ligação de seqüências reguladoras de genes a um gene repórter (um gene codificante de uma enzima tal como o gene *LacZ* na proteína verde fluorescente da água-viva), introduzindo o DNA na linhagem germinativa da mosca, e monitorando a expressão repórter na prole do embrião de moscas transgênicas (Figura 12.19). Embora as seqüências tipo selvagem 5' do gene *hunchback* sejam suficientes para ativar a expressão repórter na metade anterior do embrião, as deleções dos sítios de ligação a

Bicoid nesse elemento regulatório de ação cis reduzem ou abolem a expressão repórter (Figura 12.18b). Mais de um sítio Bicoid deve ser ocupado para gerar uma delimitação da expressão repórter, o que indica que um limiar de concentração da proteína Bicoid é necessário para ocupar vários sítios antes que a expressão gênica seja ativada. Um gene gap com poucos sítios de ligação não será ativado em locais de concentração menor.

Cada gene gap contém elementos regulatórios de ação cis com disposições diferentes de sítios de ligação, e esses sítios de ligação podem ter afinidades diferentes pela proteína Bicoid. Conseqüentemente, cada gene gap é expresso em um único domínio distinto no embrião, em resposta a níveis diferentes de Bicoid e outros gradientes de transcrição. Um tema similar é encontrado na padronização do eixo dorsoventral: os elementos regulatórios de ação cis contêm números diferentes de arranjos de sítios de ligação para Dorsal e outros fatores de transcrição dorsoventral. Conseqüentemente, são ativados genes em domínios discretos ao longo do eixo dorsoventral.

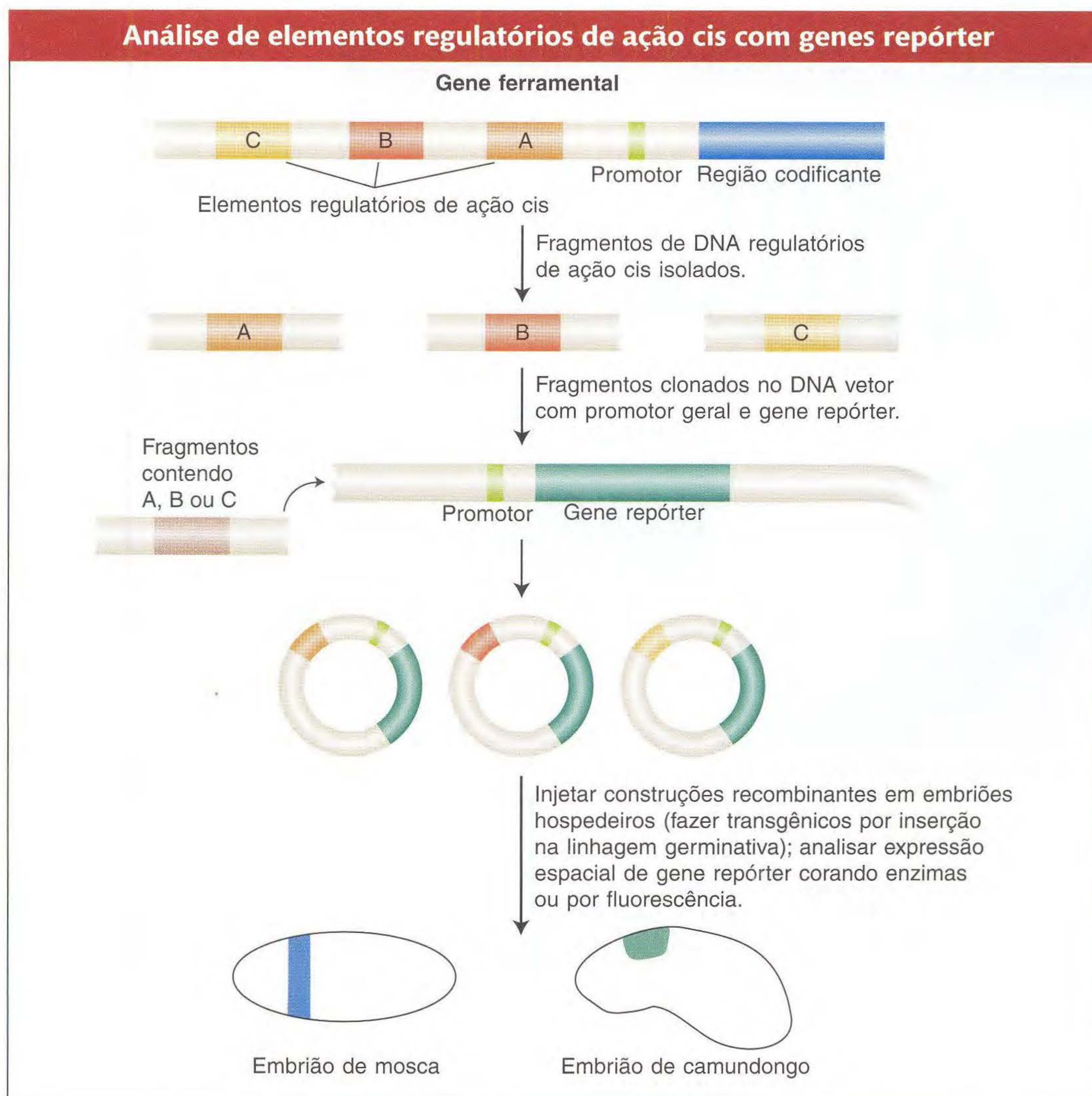


FIG. 12.19 Os loci ferramentais em geral contêm vários elementos independentes regulatórios de ação cis que controlam a expressão gênica em lugares diferentes ou em épocas diferentes durante o desenvolvimento de ambos (por exemplo, A, B, C, aqui). Esses elementos são identificados por sua habilidade, quando colocados em cis a um gene repórter e inseridos novamente em um genoma hospedeiro, para controlar o padrão, época ou nível, ou todos os três, da expressão do gene repórter. A maioria dos genes repórter codificam enzimas ou proteínas fluorescentes que podem ser facilmente visualizadas.

A resposta dependente de concentração dos genes a estímulos graduados é uma característica crucial da regulação gênica no embrião inicial de *Drosophila*. Os elementos regulatórios de ação que controlam respostas distintas contêm números diferentes e arranjos de sítios de ligação de fatores de transcrição.

Desenho das faixas: Integração dos impulsos de proteínas gap

A expressão de cada gene de regra dos pares em sete faixas é o primeiro sinal da organização periódica do embrião e futuro animal. Como tais padrões periódicos são gerados de uma informação anterior aperiódica? Antes da análise molecular da regulação do gene da regra dos pares, vários modelos foram criados para explicar a formação das faixas. Cada uma dessas idéias viu todas as sete faixas como resul-

tados idênticos em resposta a estímulos idênticos. Entretanto, o modo real pelo qual os padrões de alguns genes importantes de regra dos pares são codificados é gerado em uma faixa de cada vez. A solução para o mistério da geração de faixas destaca um dos mais importantes conceitos quanto ao controle espacial da regulação gênica em animais em desenvolvimento; isto é, os elementos regulatórios de ação cis de genes individuais são controlados independentemente.

A principal descoberta foi que cada uma das sete faixas que constituem os padrões de expressão dos genes de regra dos pares *even-skipped* e *hairy* é controlada independentemente. Considere a segunda faixa expressa pelo gene *even-skipped* (Figura 12.20a). Essa faixa fica dentro de uma região larga de expressão *hunchback* e nas bordas das regiões de expressão de duas outras proteínas gap (Figura 12.20b). Assim, dentro da área da futura faixa, existirão grandes quan-

As combinações de proteínas de efeito materno e gap controlam a formação de faixas de regra dos pares individuais

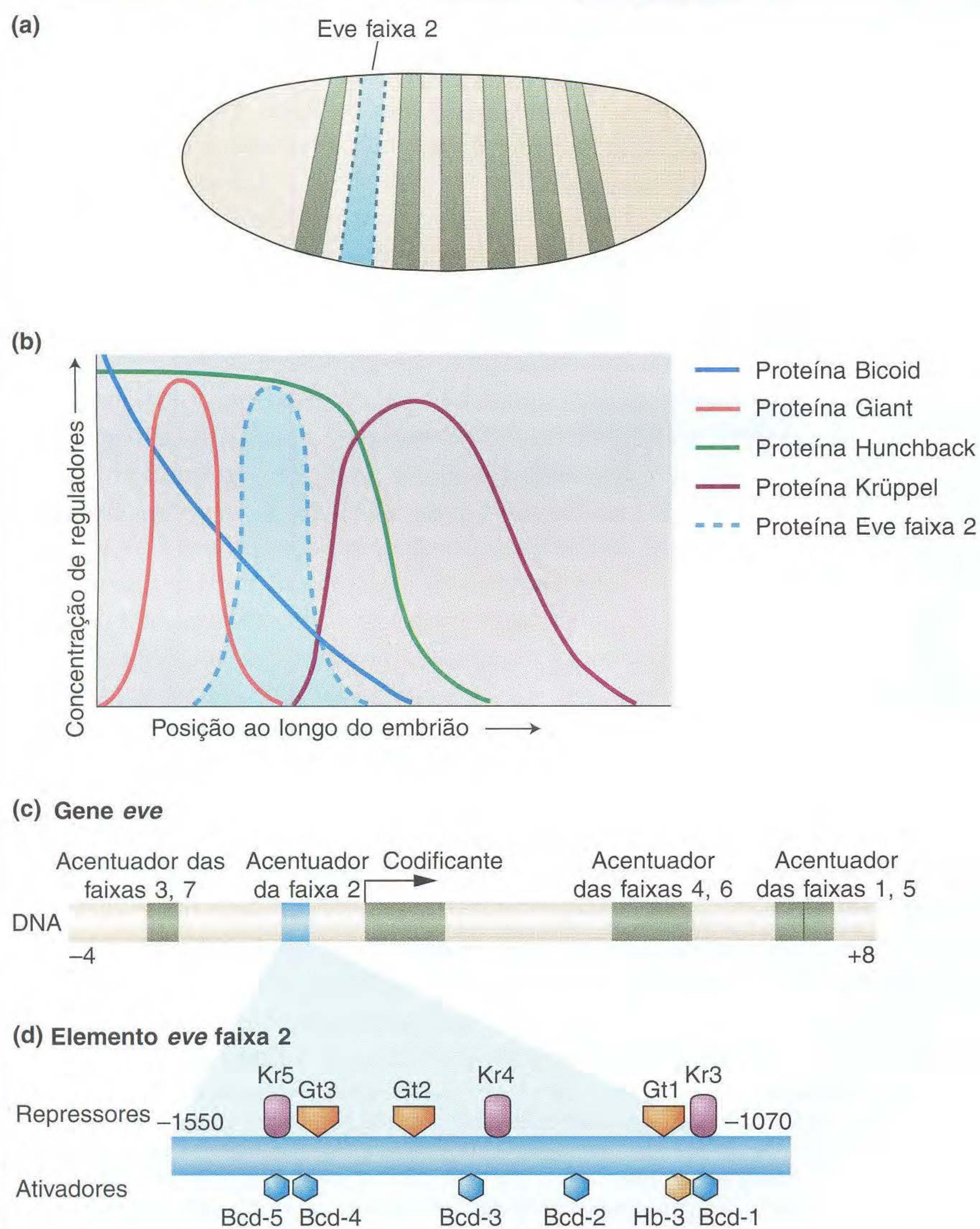


FIG. 12.20 Regulação de uma faixa de regra dos pares: controle combinatório de um elemento regulatório de ação cis independente. (a) A regulação do *eve* faixa 2 do elemento regulatório de ação cis controla a formação da segunda faixa da expressão de *eve* no embrião inicial, apenas uma das sete faixas da expressão de *eve*. (b) A faixa forma-se dentro dos domínios das proteínas Bicoid e Hunchback e na borda das proteínas gap Giant e Krüppel. Bcd e Hb são ativadores e Gt e Kr são repressores da faixa. (c) O elemento *eve* faixa 2 é apenas um dos vários elementos regulatórios de ação cis do gene *eve*, cada um dos quais controla partes diferentes da expressão *eve*. O elemento *eve* faixa 2 engloba desde cerca de 1 a 1,7 kb antecedentes à transcrição da unidade *eve*. (d) Dentro do elemento *eve* faixa 2 existem vários sítios de ligação para cada fator de transcrição (os repressores são mostrados acima do elemento; os ativadores, abaixo). O produto dessa combinação de ativadores e repressores é a expressão da estreita faixa *eve*. [De J. Gerhart and M. Kirshner, *Cells, Embryos, and Evolution*. Blackwell Science, 1997.]

tidades de proteína Hunchback e pequenas quantidades da proteína Krüppel. Haverá também uma certa concentração de proteína Bicoid de efeito materno. Nenhuma outra faixa do embrião conterá essas proteínas nessas proporções. A formação da faixa 2 é controlada por um elemento regulatório de ação cis específico, um acentuador que contém um número de sítios de ligação para essas quatro proteínas (Figura 12.20c). O acentuador da faixa 3 conterá uma combinação diferente de sítios de ligação, e assim por diante. Assim, o padrão periódico total é a soma dos diferentes conjuntos de estímulos em separados elementos regulatórios de ação cis. A análise detalhada do elemento regulatório de ação cis 2 de faixa *eve* revelou que a posição dessa “simples” faixa é controlada por não menos de quatro fatores de transcrição distribuídos aperiodicamente, incluindo uma proteína materna e três proteínas gap.

Especificamente, o elemento *eve* faixa 2 contém vários sítios para a proteína materna Bicoid, e as proteínas gap Hunchback, Giant e Krüppel (Figura 12.20d). A análise mutacional de diferentes combinações dos sítios de ligação revelou que Bicoid e Hunchback ativam a expressão do elemento *eve* faixa 2 em uma ampla região. As proteínas

Giant e Krüppel são repressores que afinam os limites da faixa apenas a algumas células de largura. O elemento *eve* faixa 2 atua, então, como um interruptor genético, integrando várias atividades de proteínas reguladoras para produzir uma faixa de três a quatro células de largura no embrião.

A regulação dos elementos regulatórios de ação cis pelas combinações de ativadores e repressores é um tema comum na regulação espacial da expressão do gene. Os padrões complexos de estímulos são geralmente integrados para produzir padrões mais simples de respostas.

Tornando os segmentos diferentes: Integração dos estímulos Hox

A atividade combinada e sequencial do efeito materno, gap, regra dos pares e proteínas de polaridade segmentar estabelece o plano corpóreo segmentado básico do embrião e larva. Como são estabelecidas as identidades segmentais diferentes pelas proteínas Hox? Esse processo tem dois aspectos. Primeiro, os genes *Hox* são expressos em domínios diferentes

Proteínas Hox reprimem a formação de apêndices no abdome

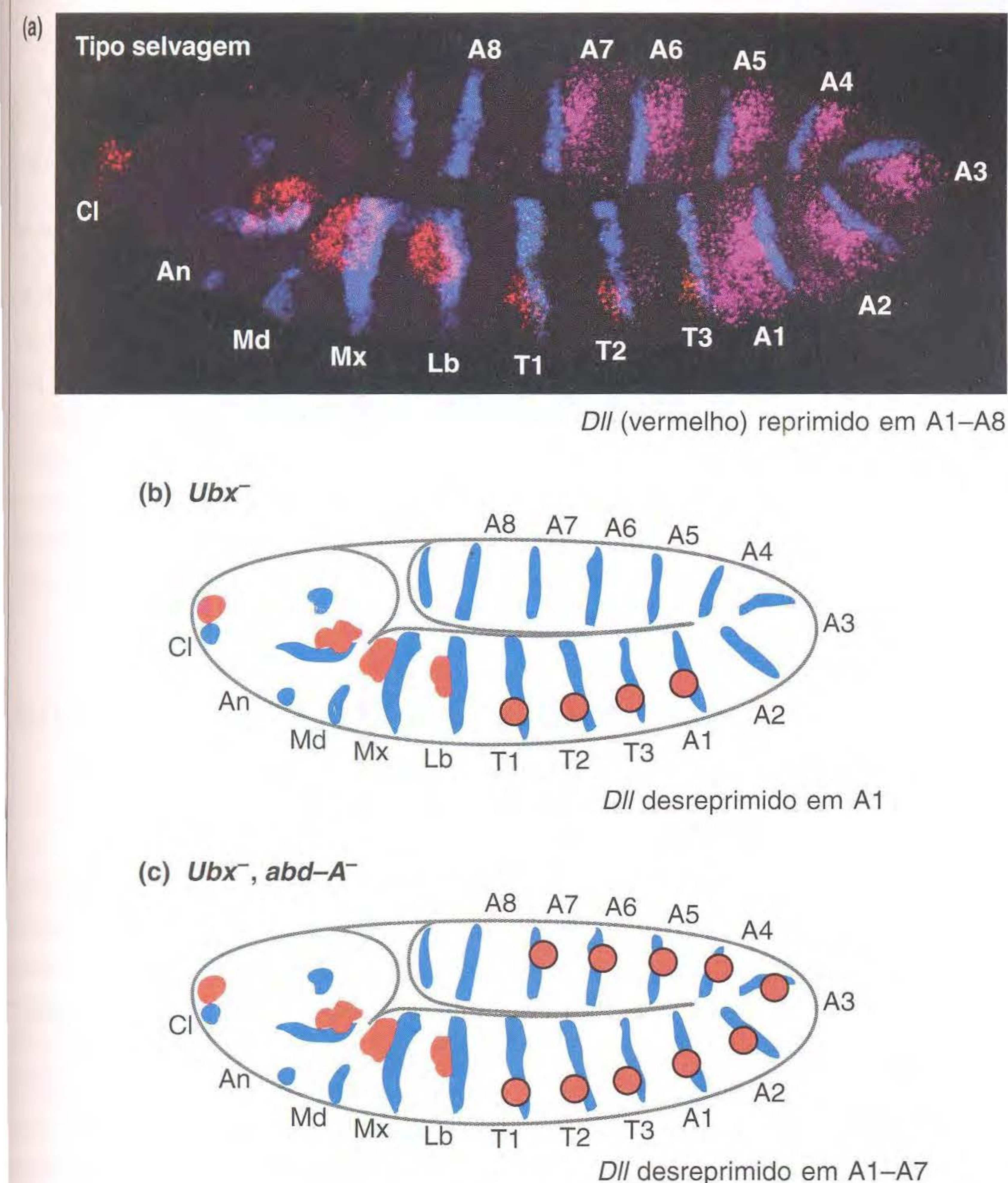


FIG. 12.21 A ausência de membros no abdome é controlada pelos genes *Hox*. (a) A expressão do gene *Distal-less* (*Dll*) (vermelho) marca a produção de apêndices futuros; expressão do gene *Hox Ultrabithorax* (púrpura) marca a posição dos segmentos abdominais A1 até A7; e a expressão do gene *engrailed* (azul) marca o posterior de cada segmento. (b) Representação esquemática do embrião *Ubx⁻* mostrando que a expressão de *Dll* (círculos vermelhos) é desreprimida no segmento A1. (c) Representação esquemática do embrião *Ubx⁻ abd-A⁻* mostrando que a expressão de *Dll* (círculos vermelhos) é desreprimida nos primeiros sete segmentos abdominais. [(a) Fotomicrografia de Dave Kosman, Ethan Bier, and Bill McGinnis; (b e c) baseada em B. Gebelein, D. J. McKay and R. S. Mann, "Direct Integration of Hox and Segmentation Gene Inputs During *Drosophila* Development," *Nature* 431, 2004, 653–659.]

ao longo do eixo ântero-posterior. A expressão do gene *Hox* é amplamente controlada por proteínas de segmentação, especialmente proteínas gap, por mecanismos que são similares aos já descritos (bem como alguma regulação cruzada por proteínas *Hox* de outros genes *Hox*). A regulação de genes *Hox* não será considerada em profundidade aqui. O segundo aspecto do controle *Hox* da identidade segmentar é a regulação dos genes-alvo por proteínas *Hox*. Examinaremos um exemplo que ilustra bem como uma característica importante do plano corpóreo da mosca é controlado através da integração de muitos estímulos por um único elemento regulatório de ação cis.

Os membros pareados, partes da boca e antenas de *Drosophila* desenvolvem-se de pequenas populações de cerca de 20 células. Estruturas diferentes desenvolvem-se de segmentos diferentes da cabeça e tórax, enquanto o abdome não tem membros. O primeiro sinal de desenvolvimento dessas estruturas é a ativação de genes reguladores dentro de pequenos grupos de células, que são chamados de *primórdios* de apêndices. A expressão do gene *Distal-less* (*Dll*) marca o início do desenvolvimento dos apêndices. Esse gene é um dos alvos importantes dos genes *Hox*, e sua função é necessária para o desenvolvimento subsequente das partes distais de cada um desses apêndices. Os pequenos grupos de células que expressam *Distal-less* surgem em vários segmentos da cabeça e em cada um dos três segmentos torácicos, mas não no abdome (Figura 12.21a).

Como a expressão *Distal-less* é restrita a segmentos mais anteriores? Reprimindo a expressão no abdome. Várias linhas de evidência revelaram que o gene *Distal-less* é reprimido pelas proteínas *Hox*, as proteínas *Ultrabithorax* e *Abdominal-A*, funcionando em colaboração com duas proteínas de segmentação. Note que a Figura 12.6 do *Ultrabithorax* é expressa em segmentos abdominais de um a sete, e *Abdominal-A* é expresso em segmentos abdominais dois a sete, superpondo-se a todos menos o primeiro segmento coberto por *Ultrabithorax*. Nos embriões mutantes *Ultrabithorax*, a expressão de *Distal-less* expande-se ao primeiro segmento abdominal (Figura 12.21b) e, nos embriões duplo-mutantes *Ultrabithorax/Abdominal-A*, a expressão de *Distal-less* estende-se através dos primeiros sete segmentos abdominais (Figura 12.21c), indicando que ambas as proteínas são necessárias para a repressão da expressão de *Distal-less* no abdome.

O elemento regulatório de ação cis responsável pela expressão de *Distal-less* no embrião foi identificado e caracterizado em detalhe (Figura 12.22a). Ele contém dois sítios de ligação para as proteínas *Hox*. Se esses dois sítios de ligação são mutados de modo que as proteínas *Hox* não possam ligar-se, a expressão de *Distal-less* é desreprimida no abdome (Figura 12.22b). Várias proteínas adicionais colaboram com as proteínas *Hox* na repressão de *Distal-less*. Duas são proteínas codificadas por genes de polaridade de segmento, *Sloppy-paired* (*Slp*) e *Engrailed* (*En*). As proteínas *Sloppy-paired* e *Engrailed* são expressas em faixas que marcam os compartimentos anterior e posterior de cada segmento, respectivamente. Cada proteína também se liga ao elemento regulatório de ação cis *Distal-less*. Quando o sítio de ligação *Sloppy-paired* é mutado no elemento regulatório de ação cis,

a expressão do gene repórter é desreprimida nos compartimentos de segmentos abdominais (Figura 12.22c). Quando o sítio de ligação de *Engrailed* está mutado, a expressão de repórter é desreprimida nos compartimentos posteriores de cada segmento abdominal (Figura 12.22d). Quando os sítios de mutação de ambas as proteínas estão mutados, a expressão do gene repórter é desreprimida em ambos os compartimentos de cada segmento abdominal, do mesmo modo que quando os sítios de ligação *Hox* estão mutados (Figura 12.22e). Duas outras proteínas, chamadas de *Extradenticle* e *Homothorax*, que são amplamente expressas em cada segmento, também se ligam ao elemento regulatório de ação cis *Distal-less* e são necessárias para a repressão transcricional no abdome (Figura 12.22f).

Assim, duas proteínas *Hox* e quatro outros fatores de transcrição ligam-se a um trecho de 57 pares de bases e atuam juntos para reprimir a expressão *Distal-less* e, assim, a formação de apêndices no abdome. A repressão da expressão *Distal-less* é uma demonstração clara de como as proteínas *Hox* regulam a identidade segmentar e o número de reiteradas estruturas corpóreas. É também uma boa ilustração de como diversos estímulos regulatórios convergem e atuam combinatoriamente em elementos regulatórios de ação cis. Nesse caso, a presença de sítios de ligação *Hox* não é suficiente para a repressão transcricional: as interações colaborativas e cooperativas são necessárias entre várias proteínas para reprimir completamente a expressão no abdome.

A regulação combinatorial e cooperativa da transcrição gênica impõe maior especificidade nos padrões espaciais de expressão gênica e permite sua maior diversidade.

Embora a diversidade evolutiva não tenha sido explicitamente abordada neste capítulo, a presença de vários elementos independentes regulatórios de ação cis para cada gene ferramental tem profundas implicações para a evolução da forma. Especificamente, a modularidade desses elementos permite mudanças em um aspecto da expressão gênica independente de outras funções gênicas. A evolução da regulação gênica tem um importante papel na evolução da morfologia. Voltaremos a esse tópico no Capítulo 19.

► 12.5 Regulação Pós-transcricional da Expressão Gênica no Desenvolvimento

Embora a regulação transcricional seja um meio importante de restringir a expressão dos produtos gênicos para definir áreas durante o desenvolvimento, não é um meio exclusivo de fazer isso. A recomposição alternativa do RNA também contribui para a regulação gênica, e, assim, a regulação da tradução do mRNA em proteínas e microRNA (miRNA). Em cada caso, as seqüências regulatórias no RNA são reconhecidas por fatores de recomposição, proteínas de ligação ao mRNA ou miRNA, e controlam a estrutura do produto

As proteínas Hox e de polaridade de segmento controlam a localização de apêndices

ELEMENTO REGULATÓRIO DE AÇÃO CIS

(a) Tipo selvagem



(b) Mutações Hox



(c) Mutação Slp



(d) Mutação En



(e) Mutações Slp, En



(f) Mutações Exd, Hth



EXPRESSÃO DE GENE REPÓRTER

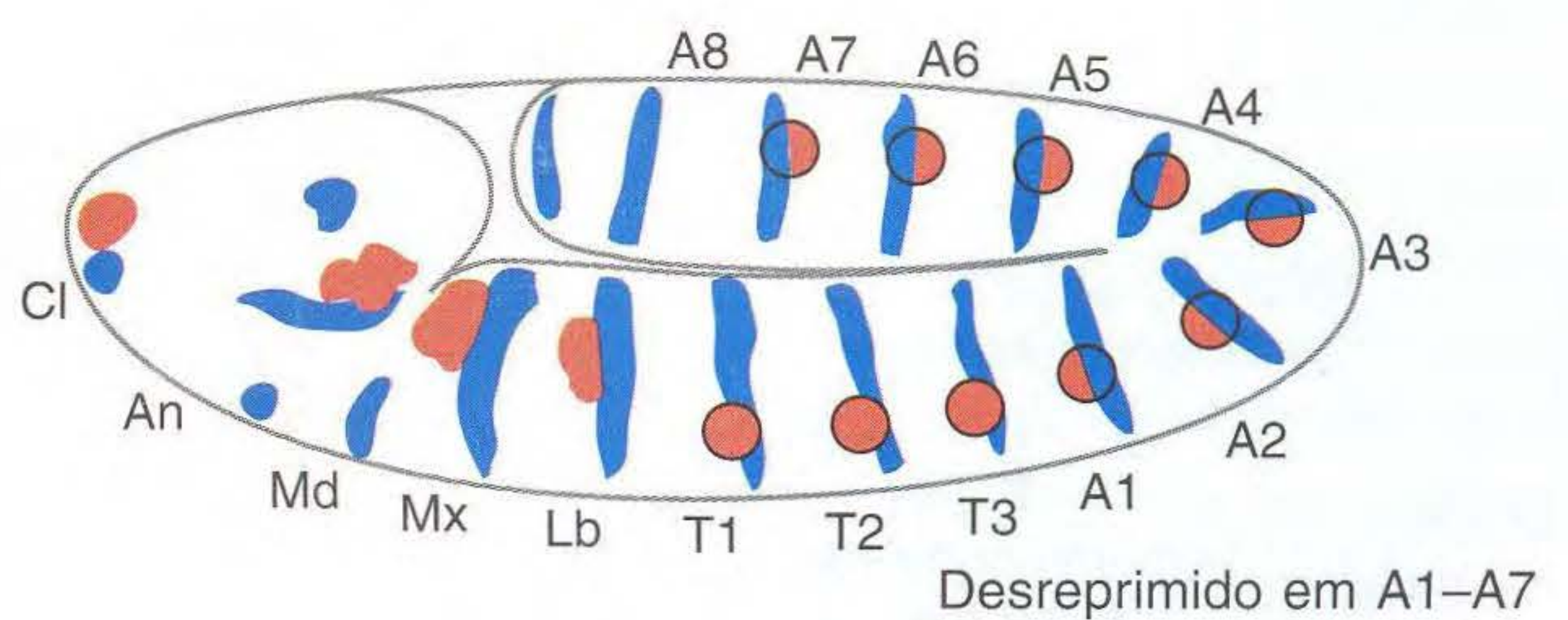
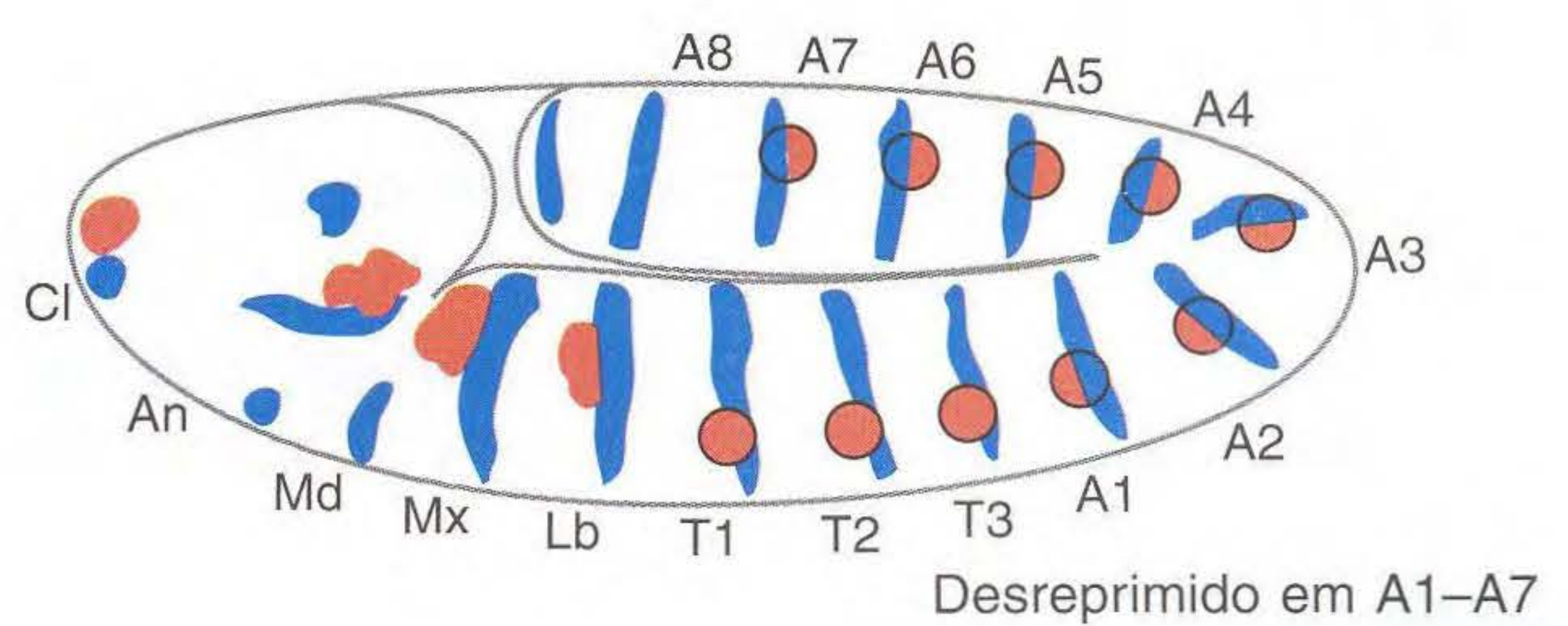
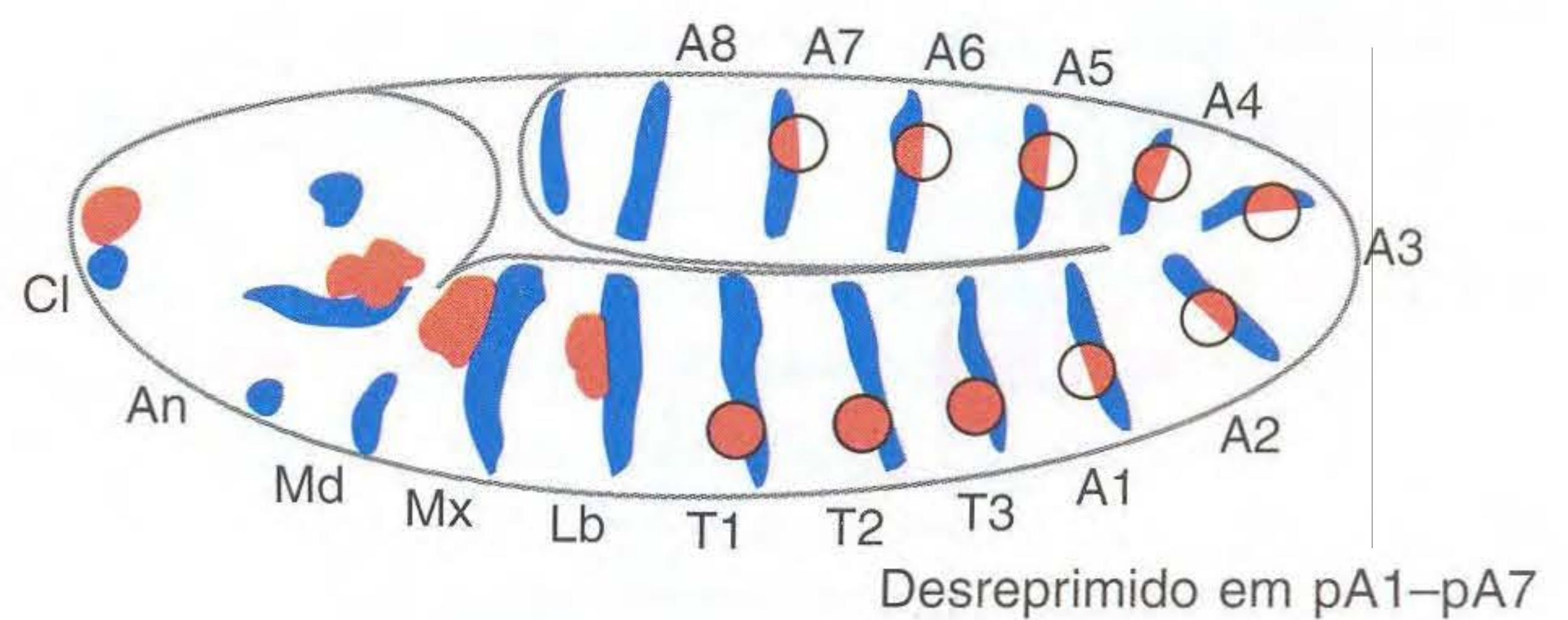
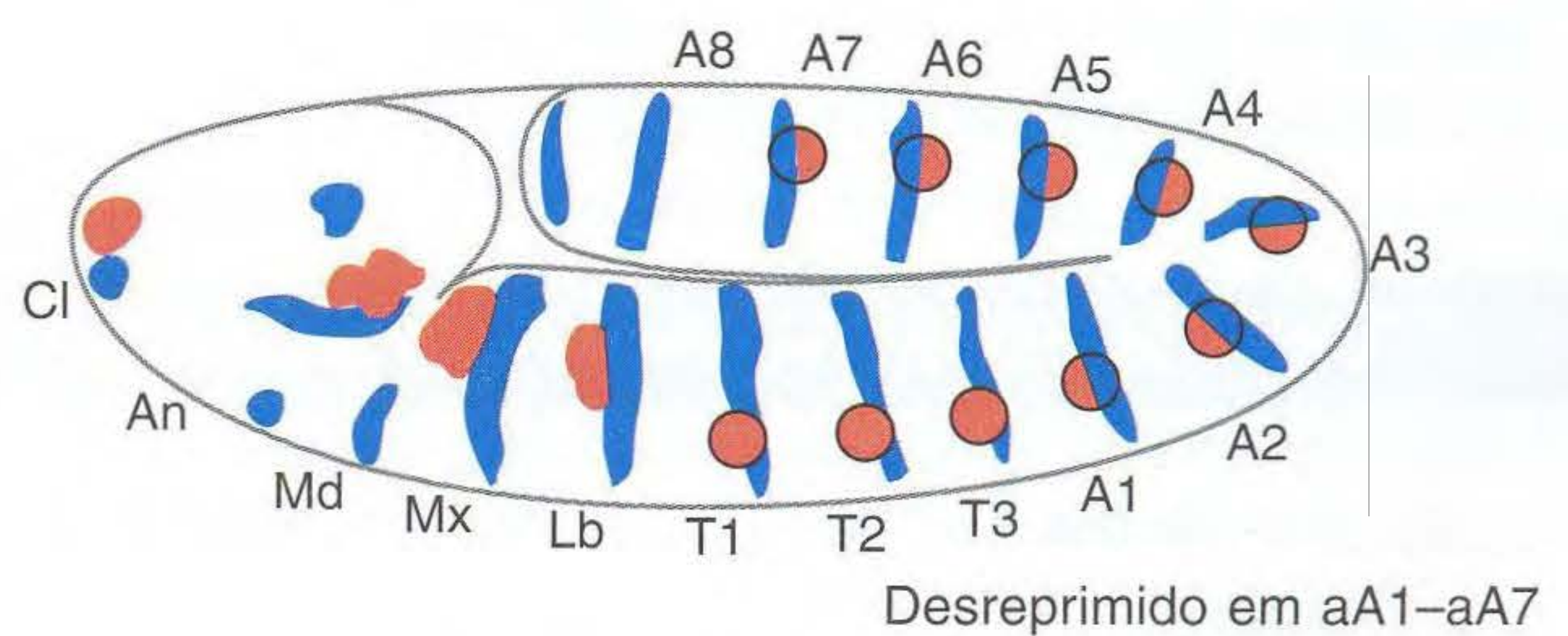
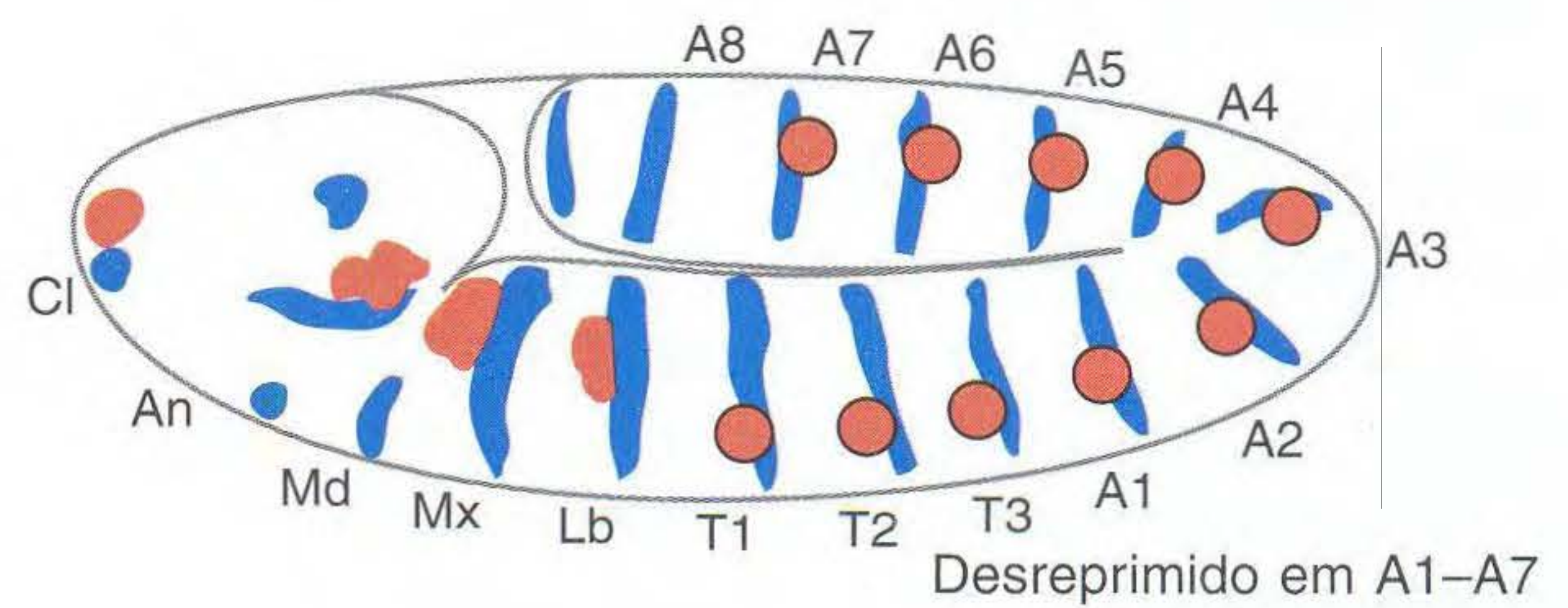
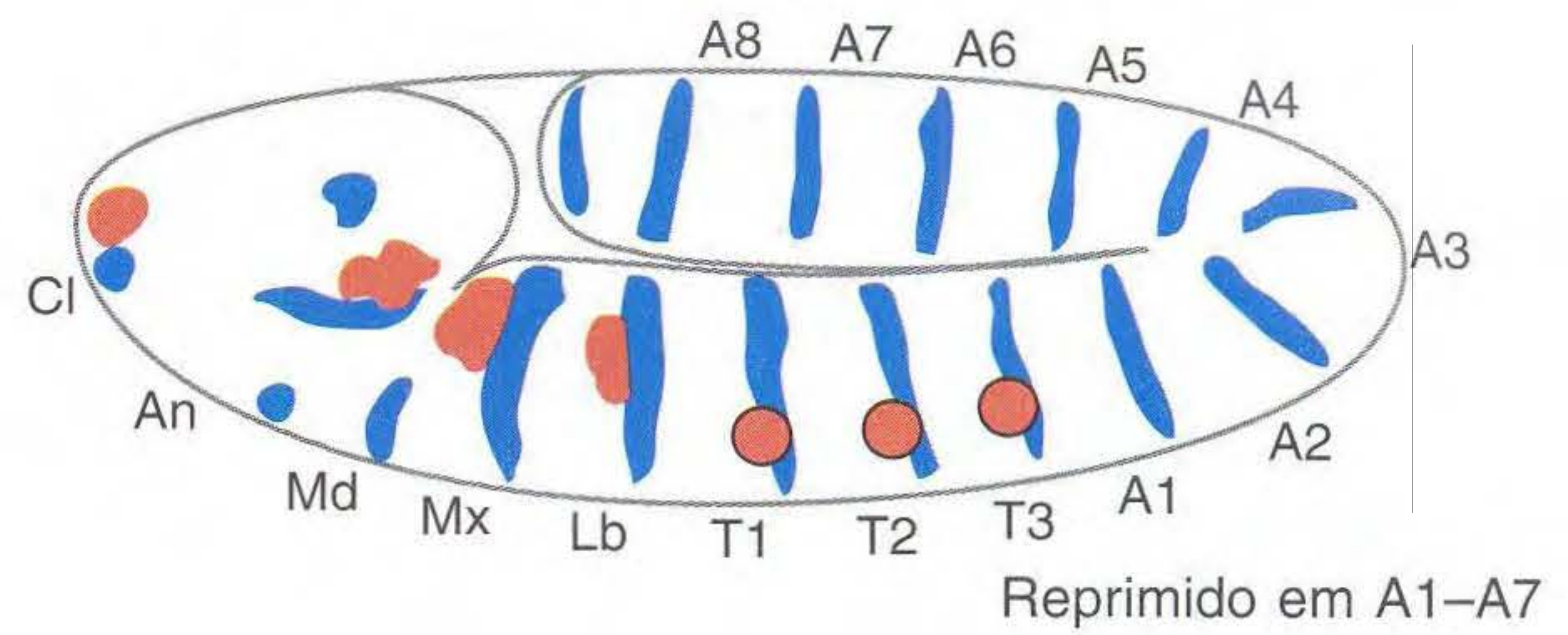


FIG. 12.22 Integração de Hox e estímulos de segmentação de proteínas por um elemento regulatório de ação cis. (a, esquerda) Um elemento regulatório de ação cis do gene *Dll* controla a expressão de *Dll* no abdome por um conjunto de fatores de transcrição. (a, direita) A expressão de *Dll* (vermelho) estende-se ao tórax, mas não ao abdome em um embrião tipo selvagem. (b-f) Mutações nos respectivos sítios de ligação mostram desrepressão de *Dll* em vários padrões no abdome. Os sítios de ligação são: Slp, Sloppy-paired; Hox1 e Hox2, Ultrabithorax e Abdominal-A; Exd, Extradenticle; En, Engrailed; Hth, Homothorax. [Baseado em dados de B. Gebelein, D. J. McKay, and R. S. Mann, "Direct Integration of Hox and Segmentation Gene Inputs During *Drosophila* Development," *Nature* 431, 2004, 653-659.]

proteico, sua quantidade ou a localização onde a proteína é produzida. Veremos um exemplo de cada tipo da interação regulativa ao nível do RNA.

Recomposição do RNA e determinação do sexo em *Drosophila*

Uma decisão desenvolvimental fundamental nos organismos de reprodução sexuada é a especificação do sexo. Nos animais, o desenvolvimento de muitos tecidos segue vias diferentes, dependendo do sexo do animal individual. Em *Drosophila*, foram identificados muitos genes que controlam a determinação do sexo pela análise de fenótipos mutantes nos quais a identidade sexual é alterada ou ambígua.

O gene *doublesex* (*dsx*) tem um papel central em controlar a identidade sexual do tecido somático (linhagem não-germinativa). As mutações nulas em *dsx* fazem com que os machos se desenvolvam como intermediários *intersexo*, que perderam as diferenças distintas entre os tecidos masculinos e femininos. Embora a função de *dsx* seja necessária em ambos os sexos, diferentes produtos gênicos são produzidos

pelo locus em sexos diferentes. Nos machos, o produto é uma isoforma específica maior, *Dsx^M*, que contém uma única região C-terminal de 150 aminoácidos não encontrada na isoforma *Dsx^F* específica das fêmeas, que, em vez disso, contém uma sequência única de 30 aminoácidos no terminal carboxila. Cada forma da proteína *Dsx* é um fator de transcrição de ligação ao DNA que, aparentemente, liga-se às mesmas sequências de DNA. Entretanto, as atividades das duas isoformas diferem: *Dsx^F* ativa alguns genes-alvo nas fêmeas que *Dsx^M* reprime nos machos.

As formas alternativas da proteína *Dsx* são geradas pela recomposição alternativa do RNA transcrito de *dsx* primário. Assim, nesse caso, a escolha dos sítios de corte deve ser regulada para produzir mRNA finais que codificam proteínas diferentes. Os vários fatores genéticos que influenciam a expressão de *Dsx* e a determinação do sexo foram identificados por mutações que afetam o fenótipo sexual.

Um regulador importante é o produto do gene *transformer* (*tra*). Enquanto as mutações em *tra* não têm efeito nos machos, as moscas fêmeas XX que têm mutações *tra* são transformadas no fenótipo masculino. A proteína *Tra* é um fator

Uma cascata de recomposição alternativa regula a determinação do sexo em *Drosophila*

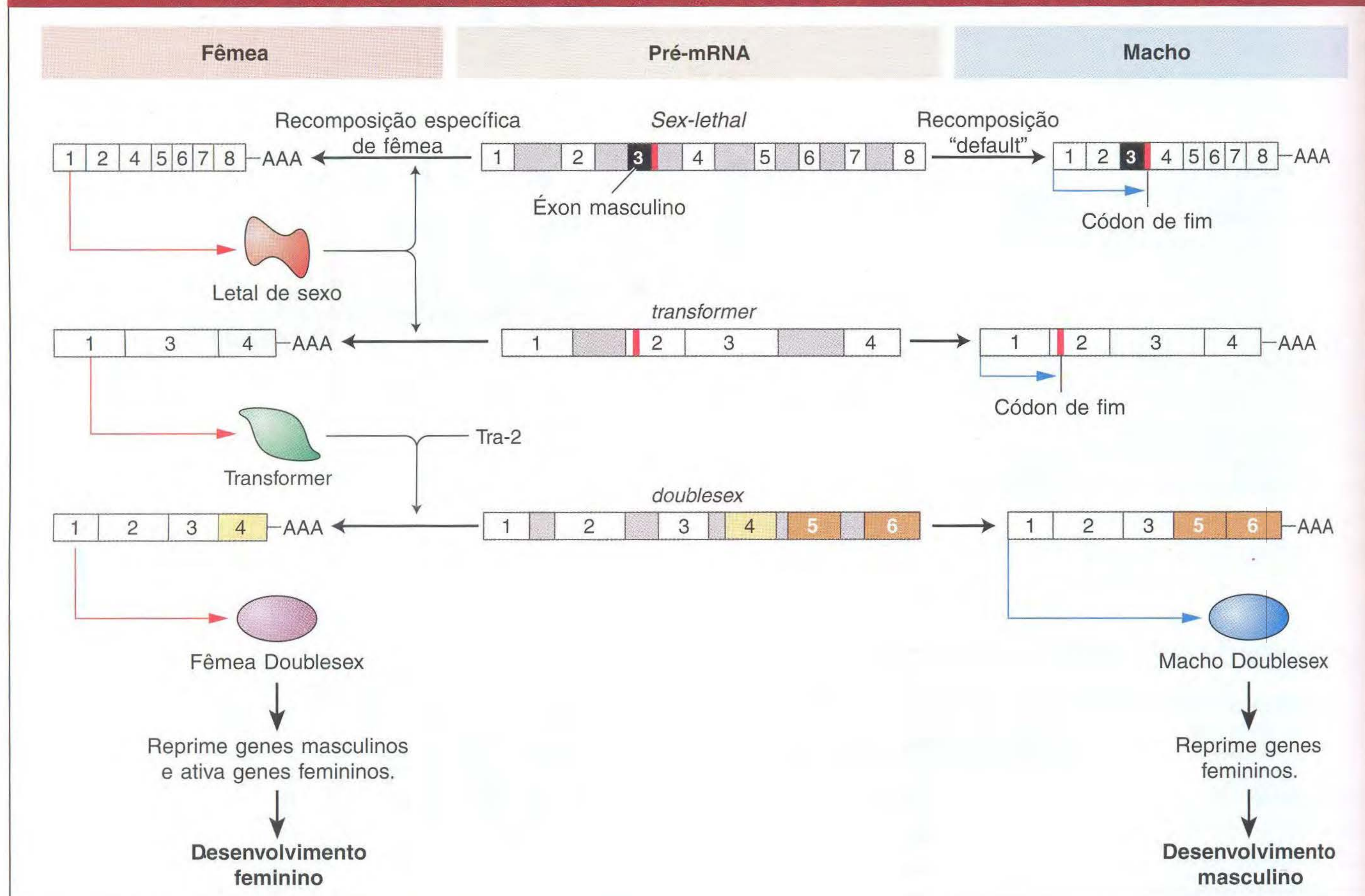


FIG. 12.23 Três pré-mRNA de genes importantes de determinação do sexo em *Drosophila* são alternativamente recompostos. A via específica das fêmeas é mostrada à esquerda e a via específica dos machos é mostrada à direita. Os pré-mRNA são idênticos em ambos os sexos e mostrados no meio. No macho, nos mRNA *sex-lethal* e *transformer*, existem códons de fim que terminam a tradução. Essas sequências são removidas por recomposição para produzir proteínas funcionais na fêmea. As proteínas Transformer e Tra-2 então recompõem o pré-mRNA de *doublesex* da fêmea para produzir a isoforma específica da fêmea da proteína *Dsx*, que difere da isoforma específica do macho pela recomposição alternativa de vários éxons. [De S. S. Gilbert, *Developmental Biology*, 7th ed. Sinauer, 2003.]

de recomposição alternativa que afeta as escolhas de corte no RNA transcrito de *dsx*. Na presença de Tra (e uma proteína correlata Tra-2), ocorre uma recomposição que incorpora o éxon 4 do gene *dsx* ao transcrito *dsx^F* final (Figura 12.23), mas não os éxons 5 e 6. Os machos não têm Tra; logo, essa recomposição não ocorre, e os éxons 5 e 6 são incorporados ao transcrito *dsx^M*, mas não ao éxon 4 (Figura 12.23).

A proteína Tra explica como formas alternativas de Dsx são expressas, mas como a expressão da própria Tra é regulada para diferir em fêmeas e machos? O próprio RNA de *tra* é recomposto alternativamente. Nas fêmeas, o fator de recomposição codificado pelo gene *Sex-lethal* (*Sxl*) está presente. Esse fator de recomposição liga-se ao RNA de *tra* e impede o evento da recomposição que, de outro modo, iria incorporar um éxon que contém um códon de fim. Nos machos, não é feita nenhuma proteína Tra, pois esse códon de fim está presente.

A produção da proteína Sex-lethal é, por sua vez, regulada tanto pela recomposição do RNA quanto por fatores que alteram o nível da transcrição. O nível da transcrição de *Sxl* é controlado por ativadores no cromossomo X e repressores nos autossomos. Nas fêmeas, a ativação de *Sxl* prevalece e a proteína *Sxl* é produzida, regulando a recomposição do RNA de *tra* e fazendo um *feedback* para regular a recomposição do próprio RNA de *Sxl*. Nas fêmeas, um códon de fim é recomposto de modo que a produção da proteína *Sxl* possa continuar. Entretanto, nos machos, onde nenhuma proteína *Sxl* está presente, o códon de fim ainda está presente no RNA transcrito não-recomposto de *Sxl*, e nenhuma proteína *Sxl* pode ser produzida.

Essa cascata de recomposição de RNA específico de sexo em *D. melanogaster* ilustra um modo pelo qual o genótipo específico de cromossomo sexual leva a formas diferentes de proteínas regulatórias que são expressas em um sexo e não no outro. Curiosamente, a regulação genética da determinação do sexo difere muito entre as espécies animais, pois o genótipo sexual pode levar à expressão diferencial de genes regulatórios por meio de vias distintamente diferentes. Entretanto, as proteínas relacionadas a Dsx têm papéis na diferenciação sexual em uma ampla gama de animais, incluindo os humanos. Assim, embora existam muitos modos para gerar a expressão diferencial de fatores de transcrição, uma família de proteínas similares parece estar subjacente a grande parte da diferenciação sexual.

Regulação da tradução do mRNA e linhagem celular em *C. elegans*

Em muitas espécies animais, o desenvolvimento do embrião envolve a repartição de células em grupos de células em linhagens discretas que darão origem a tecidos distintos no adulto. Esse processo é mais bem compreendido no verme nematódeo *C. elegans*, no qual o animal adulto é composto apenas de cerca de 1.000 células somáticas (um terço das quais é representado por células nervosas) e um número similar de células germinativas na gônada. A construção simples, o rápido ciclo de vida e a transparência de *C. elegans* o tornaram um poderoso modelo para a análise desen-

volvimental (veja o boxe Organismo-modelo *Caenorhabditis elegans*, adiante). Todas essas linhagens de células animais foram mapeadas em uma série de estudos elegantes liderados por John Sulston no Medical Research Council (MRC) Laboratory em Cambridge, Inglaterra. Sistemáticas triagens genéticas quanto a mutações que perturbam ou estendem linhagens celulares forneceram muitas informações sobre o controle genético das decisões das linhagens. A genética de *C. elegans* foi especialmente importante na compreensão do papel da regulação pós-transcricional no RNA, e examinaremos aqui dois mecanismos: (1) controle da tradução por proteínas de ligação ao mRNA e (2) controle de miRNA da expressão gênica.

Controle traducional no embrião inicial

Primeiro veremos como começa uma linhagem celular. Após duas divisões celulares, o embrião de *C. elegans* contém quatro células, chamadas de blastômeros. Cada célula começará uma linhagem distinta, e as descendentes das linhagens separadas terão destinos diferentes. Já nesse estágio, são observadas diferenças nas proteínas presentes nos quatro blastômeros. Pelo que já aprendemos, muitas dessas proteínas são ferramentas que determinam quais genes serão expressos nas células descendentes. O que é surpreendente é que os mRNA que codificam algumas das proteínas ferramentas dos vermes estão presentes em *todas* as células do embrião inicial. Entretanto, em uma célula específica, apenas alguns desses mRNA serão traduzidos em proteínas. Assim, no embrião de *C. elegans*, a regulação pós-transcricional é crítica para a especificação apropriada dos destinos celulares iniciais. Durante a primeira divisão celular, a polaridade dentro do zigoto leva à repartição de moléculas reguladoras para células embrionárias específicas. Por exemplo, o gene *glp-1* codifica uma proteína receptora transmembranar (relacionada ao receptor Notch de moscas e outros animais). Embora o mRNA de *glp-1* esteja presente em todas as células do estágio de quatro células, a proteína GLP-1 é traduzida apenas nas duas células anteriores ABa e ABp (Figura 12.24a). Essa expressão localizada de GLP-1 é crítica para o estabelecimento de destinos distintos. As mutações que abolem a função *glp-1* no estágio de quatro células altera os destinos das descendentes ABp e ABa.

GLP-1 está localizada nas células anteriores, reprimindo sua tradução nas células posteriores. A repressão da tradução de GLP-1 requer seqüências na 3' UTR do mRNA de *glp-1* especificamente, uma região de 61 nucleotídeos chamada de região de controle espacial (SCR). A importância de SCR foi demonstrada ligando o mRNA transcrito de genes repórteres para variantes diferentes de SCR. A deleção dessa região ou mutação de sítios importantes dentro dela faz com que o gene repórter seja expresso em todos os quatro blastômeros do embrião inicial (Figura 12.24b).

Com base em como vimos o controle transcricional, podemos supor que uma proteína liga-se a SCR para reprimir a tradução do mRNA de *glp-1*. Para identificar essas proteínas repressoras, os pesquisadores isolaram proteínas que se ligam a SCR. Uma proteína, GLD-1, liga-se especificamente a



Organismo-modelo *Caenorhabditis elegans*

O nematódeo *Caenorhabditis elegans* como Modelo para Decisões do Destino de Linhagens Celulares

Nos últimos 20 anos, os estudos do verme nematódeo *Caenorhabditis elegans* (veja Diagrama 1) avançou muito nossa compreensão sobre o controle genético das decisões de linhagens celulares. A transparência e simples constru-

ção desse animal levaram Sydney Brenner a avançar seu uso como um organismo-modelo. O verme adulto contém cerca de 1.000 células somáticas, e os pesquisadores, liderados por John Sulston, cuidadosamente mapearam toda a série de decisões de células somáticas que produzem o animal adulto.

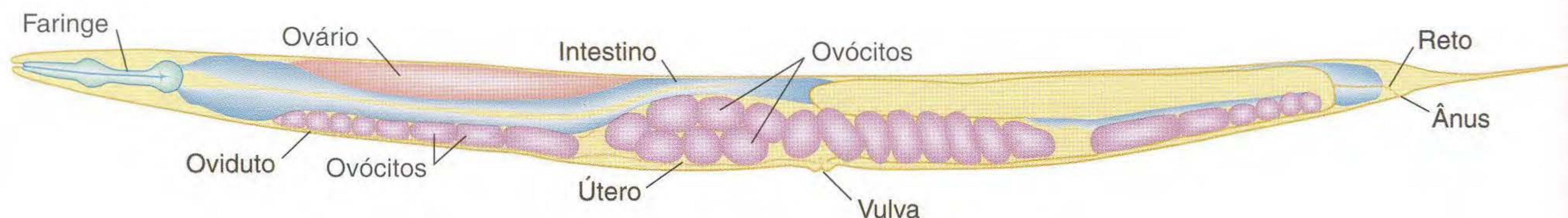
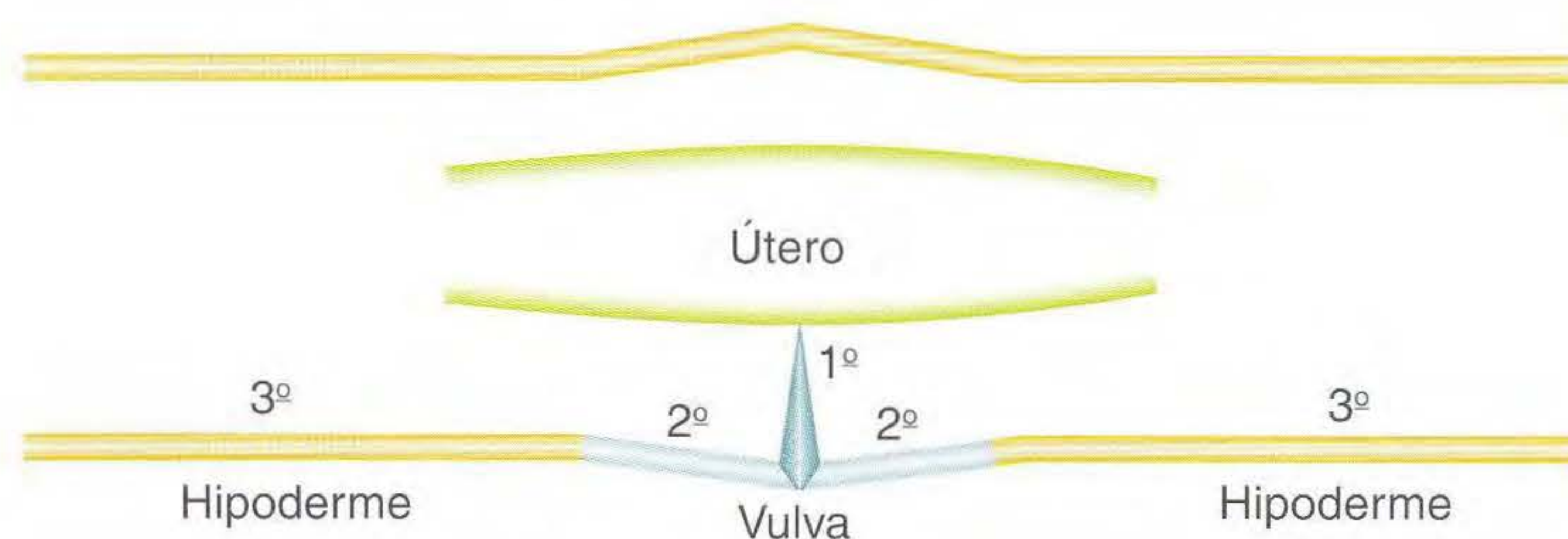


Diagrama 1. Um *Caenorhabditis elegans* adulto hermafrodita, mostrando vários órgãos.

Algumas das decisões de linhagens, tais como a formação da vulva, foram modelos-chave para as chamadas *interações indutivas* no desenvolvimento, onde a sinalização entre as células induz as mudanças do destino celular e

formação de órgãos (veja Diagrama 2). Exaustivas triagens genéticas identificaram muitos componentes participantes na sinalização e transdução de sinal na formação da vulva.

(a) Tecido derivado da 1ª, 2ª e 3ª células



(b) Pedigrees das células

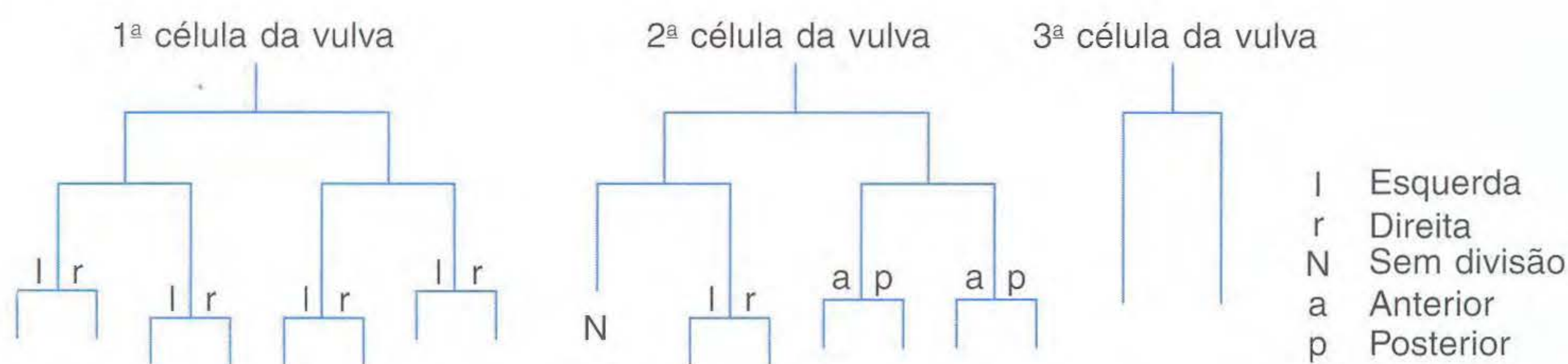


Diagrama 2. Produção das linhagens celulares vulvais. (a) As partes da anatomia vulval que são ocupadas pelas chamadas células primárias (1ª), secundárias (2ª) e terciárias (3ª). (b) As linhagens ou pedigrees das células primárias, secundárias e terciárias diferem por seus padrões de divisão celular.

Para algumas das divisões celulares embrionárias e larvais, particularmente as que contribuem para o sistema nervoso do verme, uma célula genitora dá origem a duas células genitoras, uma das quais então sofre morte celular programada. A análise de mutantes nas quais a morte celular programada é aberrante, feita por Robert Horvitz,

revelou muitos componentes das vias de morte celular programada comuns à maioria dos animais. Sydney Brenner, John Sulston e Robert Horvitz dividiram o Prêmio Nobel de 2002 em Fisiologia ou Medicina por seu trabalho pioneiro baseado em *C. elegans*.

Proteínas de ligação ao mRNA reprimem a tradução do mRNA para determinar as linhagens celulares

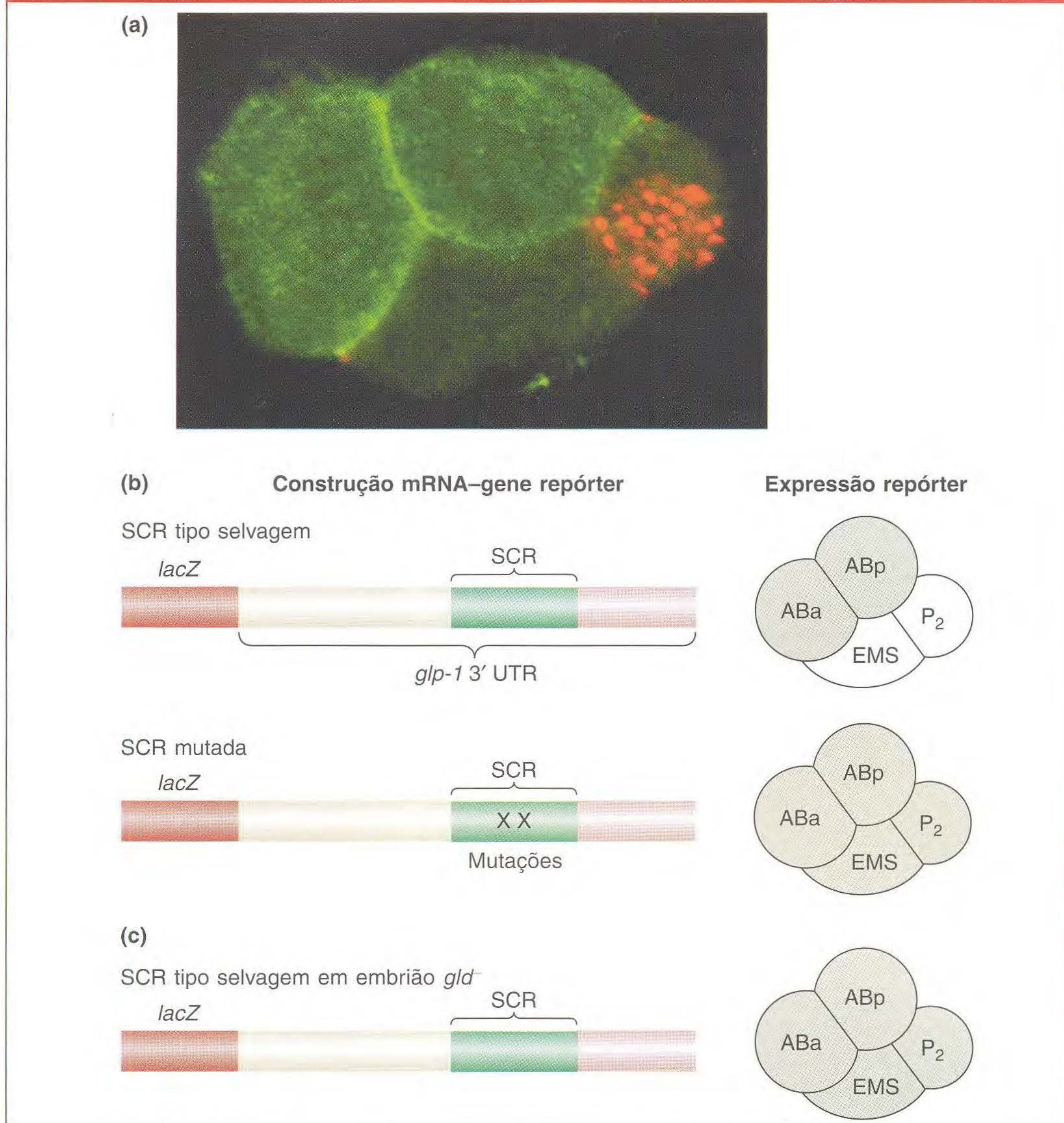


FIG. 12.24 Regulação traducional e decisões de linhagem celular no embrião inicial de *C. elegans*. (a) No estágio de quatro células do embrião de *C. elegans*, a proteína GLP-1 é expressa em duas células anteriores (verde-claro), mas não em outras células. A tradução do mRNA de *glp-1* é regulada pela proteína GLD-1 nas células posteriores. (b) Fusão de *glp-1* 3' UTR ao gene repórter *lacZ* leva à expressão repórter nas células ABa e ABp do estágio de quatro células do embrião de *C. elegans* (sombreado, à direita). As mutações em sítios de ligação GLD-1 na região de controle espacial (SCR) causam desrepressão e tradução nas linhagens EMS e P₂, como (c) perda da função de *glp-1*. [(a) Cortesia de Thomas Evans, University of Colorado Health Sciences Center.]

uma região de SCR. Além disso, a proteína GLD-1 é enriquecida em blastômeros posteriores, onde a expressão de *glp-1* é reprimida. Finalmente, quando a expressão de GLD-1 é inibida usando RNA de interferência, a proteína GLP-1 é expressa nos blastômeros posteriores (Figura 12.24c). Essa evidência sugere que GLD-1 é uma proteína repressora traducional que controla a expressão de *glp-1*.

A regulação espacial da tradução de GLP-1 é um exemplo do controle traducional do desenvolvimento ou por GLD-1. Muitos outros mRNA são regulados traducionalmente, e GLD-1 liga-se a outros mRNA-alvo nas células embrionárias e células da linhagem germinativa.

Proteínas de ligação a seqüências específicas de RNA atuam por meio de seqüências de ação cis ao RNA para regular o padrão espacial da tradução da proteína.

miRNA controla a época desenvolvimental em *C. elegans* e outras espécies

O desenvolvimento é um processo ordenado tanto temporal quanto espacialmente. Quando os eventos ocorrem é tão importante quanto onde. As mutações nos genes *heterocrônicos* de *C. elegans* foram fontes de *insight* do controle do tempo desenvolvimental. As mutações nesses genes alteram a época

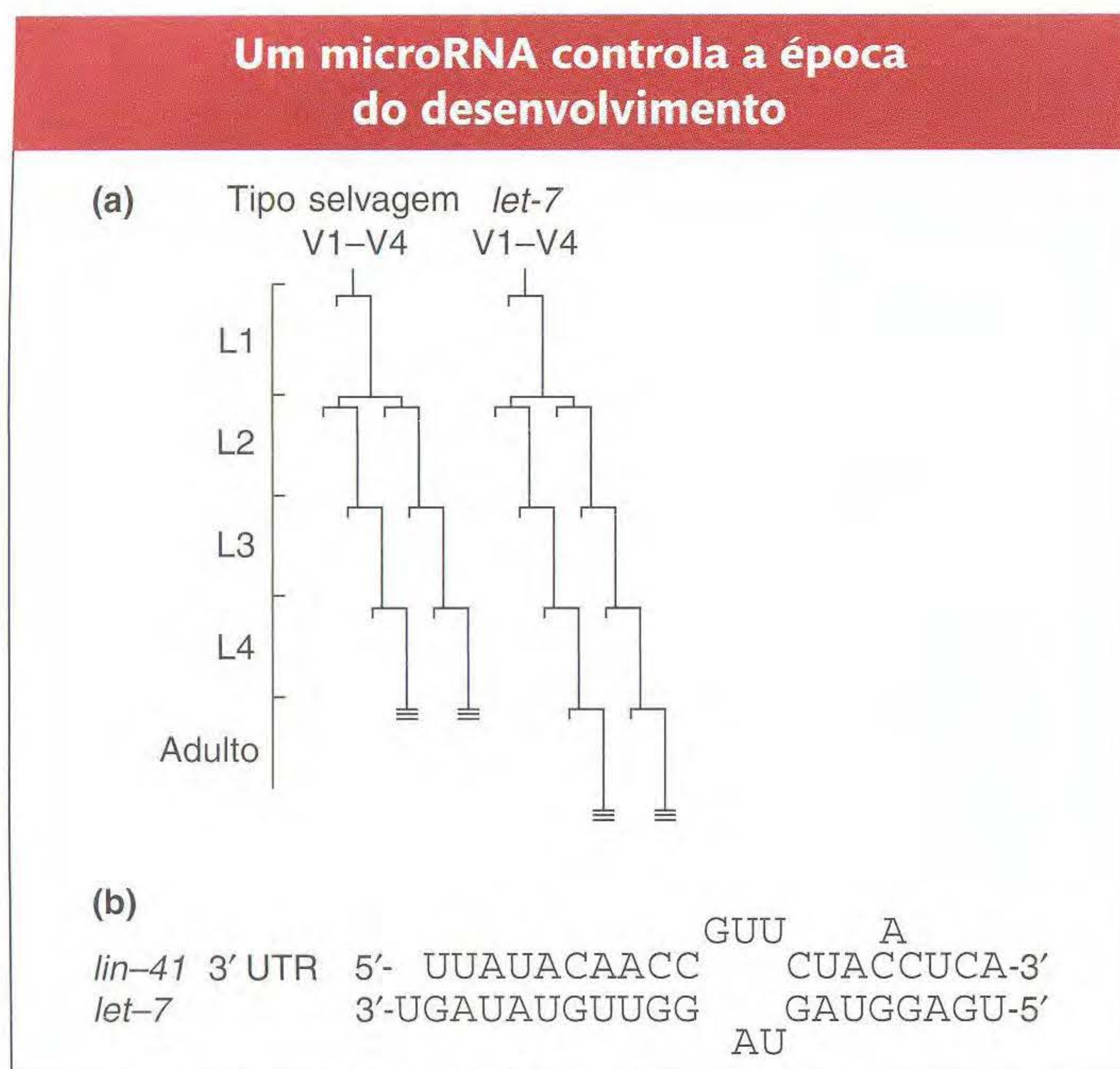


FIG. 12.25 (a) Nos mutantes *let-7*, a transição do estágio larval L4 para o adulto é retardada e as linhagens celulares das células hipodérmicas laterais (V) são reiteradas. (b) *Let-7* codifica um miRNA que é complementar a seqüências no mRNA 3' UTR da *lin-41*.

dos eventos na especificação do destino celular, fazendo com que esses eventos sejam reiterados ou omitidos. A investigação detalhada dos produtos de genes heterocrônicos levou à descoberta de um mecanismo inteiramente inesperado da regulação da expressão gênica, por meio de microRNA.

Os primeiros membros dessa classe de moléculas regulatórias a serem descobertos em *C. elegans* são RNA produzidos pelos genes *lin-4* e *let-7*. O gene *lin-4* controla a transição do primeiro para o segundo estágio larval; *let-7* regula a transição dos destinos celulares larvais adiantados para o adulto. Nos mutantes *let-7*, por exemplo os destinos das células larvais são reiterados no estágio adulto (Figura 12.25a). Contrariamente, o aumento de dosagem do gene *let-7* causa uma especificação precoce dos destinos adultos nos estágios larvares.

Nem *let-7* nem *lin-4* codificam proteínas. *Let-7* codifica um RNA final com 22 nucleotídeos regulado temporalmente e processado a partir de um precursor de aproximadamente 70 nucleotídeos. O RNA final é complementar a seqüências nas regiões 3' não traduzidas de uma variedade de genes regulados desenvolvimentalmente, e a ligação de miRNA a essas seqüências impede a tradução desses transcritos gênicos. Um desses genes-alvo, *lin-41*, também afeta a transição de larva para adulto. Os mutantes *lin-41* causam a especificação precoce dos destinos celulares dos adultos, sugerindo que o efeito da hiperexpressão de *let-7* é devido, pelo menos em parte, a um efeito da expressão de *lin-41*. O mRNA de *let-7* liga-se ao RNA de *lin-41* *in vitro* em vários sítios complementares imperfeitos (Figura 12.25b).

O papel de miRNAs no desenvolvimento de *C. elegans* estende-se bem além desses dois genes. Várias centenas de miRNA foram identificadas, e muitos genes-alvo foram

demonstrados regulados pelo miRNA. Além disso, a descoberta dessa classe de RNA reguladores levou a pesquisas de tais genes em outros genomas, e, em geral, centenas de genes miRNA candidatos foram detectados em genomas animais, inclusive os de humanos.

Surpreendentemente, o gene de miRNA *let-7* é amplamente conservado e encontrado em *Drosophila*, ascídios, moluscos, anelídeos e genomas de vertebrados (inclusive humano). O gene *lin-41* também é conservado, e as evidências sugerem que as interações regulatórias *let-7-lin41* também controlam os eventos temporais no desenvolvimento em outras espécies.

A descoberta da regulação do miRNA nos genes de desenvolvimento e o escopo do repertório de miRNA são muito recentes. Os geneticistas e outros biólogos ficaram muito entusiasmados quanto aos papéis dessa classe de moléculas reguladoras no desenvolvimento e fisiologia, levando a uma área intensa e rápida de novas pesquisas.

► 12.6 Os Vários Papéis dos Genes Ferramentais Individuais

Vimos que as proteínas ferramentais e RNA reguladores têm múltiplos papéis no desenvolvimento. Por exemplo, lembre que, na mosca, a proteína Ultrabithorax reprime a formação de membros no abdome e promove o desenvolvimento das asas posteriores no tórax. Similarmente, Sloppy-paired e Engrailed participam da geração da organização segmentar básica do embrião e colaboram com as proteínas Hox para suprimir a formação de membros. Esses papéis são apenas alguns dos muitos desempenhados por esses genes ferramentais em todo o curso do desenvolvimento da mosca. A maioria dos genes ferramentais funcionam em mais de uma época e lugar, e a maioria pode influenciar a formação ou padrões de muitas estruturas diferentes que são formadas em partes diferentes do corpo da larva ou adulto. A função de uma proteína ferramental individual (ou RNA) é quase sempre dependente do contexto, daí o motivo de a analogia com a caixa de ferramentas ser tão apropriada. Como em uma caixa de ferramentas de um carpinteiro, um conjunto comum de ferramentas pode ser usado para fazer muitas estruturas.

Para ilustrar mais vividamente esse princípio, veremos o papel de uma proteína ferramental no desenvolvimento de muitas características dos vertebrados, incluindo características presentes nos humanos. Essa proteína ferramental é o homólogo nos vertebrados do gene *hedgehog* de *Drosophila*. O gene *hedgehog* foi primeiro identificado por Nüsslein-Volhard e Wieschaus como um gene de polaridade segmentar. Ele foi caracterizado como codificando uma proteína sinalizadora secretada por células em *Drosophila*.

De moscas a dedos, penas e placas do assoalho

À medida que aumentaram as evidências de que os genes ferramentais são comuns em diferentes ramos de animais, a descoberta e caracterização dos genes ferramentais da mosca, tais como o *hedgehog*, tornaram-se uma base comum para a caracterização de genes em outro grupos, particularmente os vertebrados. A clonagem de genes homólogos baseada

O gene ferramental *Sonic hedgehog* tem múltiplos papéis

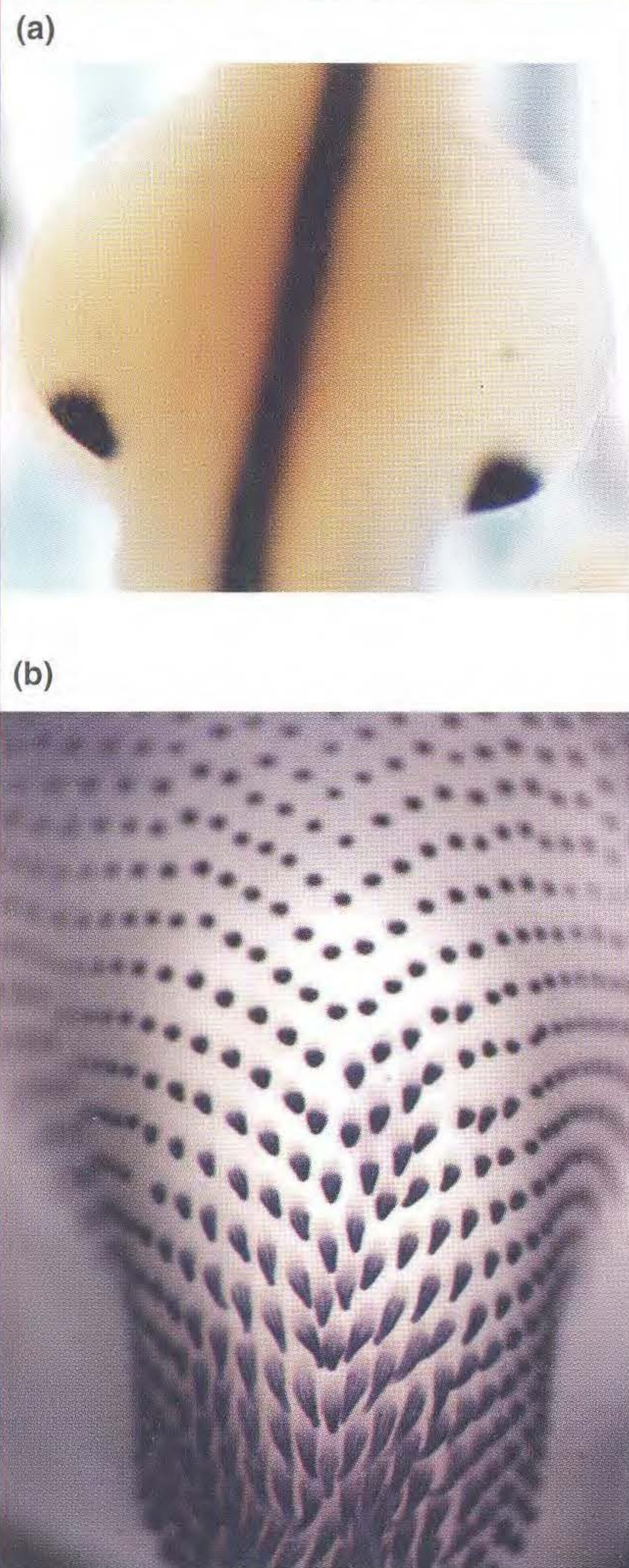


FIG. 12.26 O gene *Shh* é expresso em muitas partes diferentes do embrião de galinha em desenvolvimento (indicadas pela coloração púrpura), incluindo (a) a zona de atividade polarizante em cada um dos dois brotos dos membros em desenvolvimento e o longo tubo neural, e (b) os brotos de desenvolvimento das penas. O mRNA de *Shh* é visualizado por hibridização *in situ*. [Fotomicrografia por cortesia de (a) Cliff Tabin e (b) Matthew Harris and John Fallon.]

na similaridade de seqüências (veja o Capítulo 13) foi uma via rápida para a identificação dos genes ferramentais dos vertebrados. A aplicação dessa estratégia ao gene *hedgehog* ilustra o poder e os benefícios de usar uma homologia para descobrir genes importantes. Vários homólogos distintos de *hedgehog* foram isolados do peixe-zebra, camundongos, galinhas e humanos. No espírito da nomenclatura do gene de *Drosophila*, os três homólogos de vertebrados foram denominados *Sonic hedgehog* (personagem de videogame), *Indian hedgehog* e *Desert hedgehog*.

Um dos primeiros meios de caracterização dos papéis potenciais desses genes no desenvolvimento foi examinar onde eles são expressos. *Sonic hedgehog* (*Shh*) foi visto expres-

so em várias partes do desenvolvimento de galinhas e outros vertebrados. Mais curiosa foi sua expressão na parte posterior do desenvolvimento do broto dos membros (Figura 12.26a). Essa parte do broto dos membros foi conhecida por décadas como sendo a *zona de polarização de atividade* (ZPA), porque é um organizador responsável por estabelecer a polaridade ântero-posterior do membro e seus dedos (veja Figura 12.2b). Para testar se *Shh* pode ter um papel na função de ZPA, Cliff Tabin *et al.*, na Harvard Medical School, fizeram com que a proteína *Shh* fosse expressa na região *anterior* dos brotos dos membros de galinha em desenvolvimento. Eles observaram o mesmo efeito que o transplante de ZPA, a indução de dedos extras com polaridade reversa. Seus resultados foram uma incrível evidência de que *Shh* era o morfógeno há muito procurado produzido por ZPA.

Shh também é expresso em outros curiosos padrões em galinha e outros vertebrados. Por exemplo, *Shh* é expresso nos brotos em desenvolvimento das penas, onde tem o papel de estabelecer o padrão e polaridade da formação de penas (Figura 12.26b). *Shh* também é expresso no tubo neural em desenvolvimento dos embriões de vertebrados, em uma região chamada de *placa do assoalho* (Figura 12.26a). Os experimentos subsequentes mostraram que a sinalização *Shh* dessas células da placa do assoalho é crítica para a subdivisão dos hemisférios cerebrais e subdivisão do desenvolvimento do olho nos lados esquerdo e direito. Quando a função do gene *Shh* é eliminada por mutação no camundongo, esses hemisférios e regiões do olho não se separam, e o embrião resultante é *ciclópico*, com um olho central e um único prosencéfalo (ele também não tem membros).

Os marcantes e diversos papéis de *Shh* são um exemplo bem ilustrativo dos diferentes papéis desempenhados pelos genes ferramentais em diferentes locais e épocas no desenvolvimento. Os resultados da sinalização *Shh* são diferentes em cada caso: a via de sinalização *Shh* irá induzir a expressão de um grupo de genes no membro em desenvolvimento, um conjunto diferente no broto das penas e, ainda, outro na placa do assoalho. Como tipos diferentes de células e tecidos são capazes de responder diferentemente à mesma molécula sinalizadora? O resultado da sinalização *Shh* depende do contexto fornecido por outros genes ferramentais que estão agindo ao mesmo tempo.

A maioria dos genes ferramentais tem múltiplos papéis em tecidos e tipos de células diferentes. A especificidade de sua ação é determinada pelo contexto dado por outros genes ferramentais que atuam em combinação com eles.

► 12.7 Desenvolvimento e Doença

A descoberta das ferramentas para o desenvolvimento de moscas, vertebrados e humanos também teve um profundo efeito no estudo das bases genéticas das doenças humanas, particularmente defeitos de nascimento e câncer. Um grande número de mutações de genes ferramentais foram identificadas afetando o desenvolvimento humano e a saúde. Enfocaremos aqui alguns poucos exemplos que ilustram como a compreensão do funcionamento gênico e regulação

Polidactilia em humanos

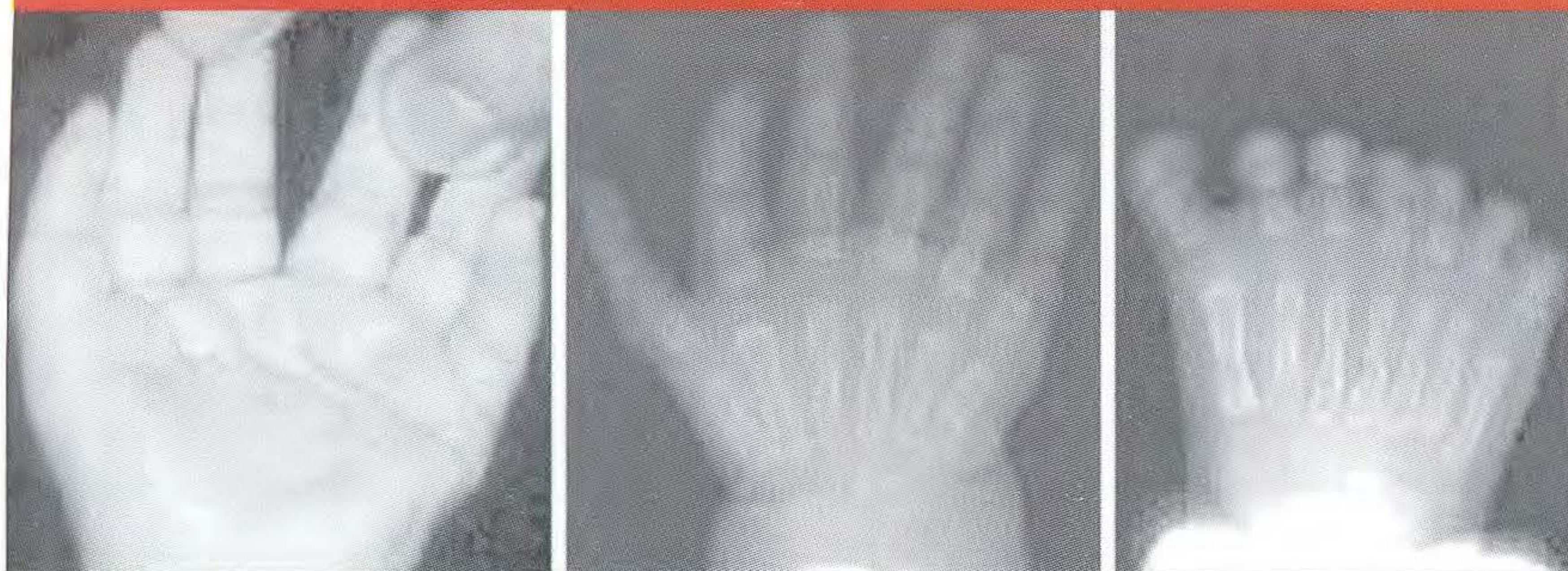


FIG. 12.27 Essa pessoa tem seis dedos em cada mão e sete artelhos em cada pé devido a uma mutação reguladora no gene *Sonic hedgehog*. [Foto por cortesia do Robert Hill, MRC Human Genetics Unit, Edinburgh, Scotland; de L. A. Lettice et al., "Disruption of a Long-Range Cis-Acting Regulator for *Shh* Causes Preaxial Polydactyly," *Proc. Natl. Acad. Sci. USA* 99, 2002, 7548.]

em modelos animais foi traduzida em uma melhor compreensão da biologia humana.

Polidactilia

Uma síndrome relativamente comum em humanos é o desenvolvimento de dedos extras, parciais ou completos, nas mãos e pés. Essa condição, chamada de *polidactilia*, surge em cerca de 5 a 17 de cada 10.000 nativos. Nos casos mais marcantes, a condição está presente tanto nas mãos quanto nos pés (Figura 12.27). A polidactilia ocorre amplamente nos vertebrados, em gatos, galinhas, camundongos e outras espécies.

A descoberta do papel de *Shh* nos padrões dos dedos levou os geneticistas a investigar se o gene *Shh* foi alterado em humanos polidáctilos e em outras espécies. De fato, algumas mutações de polidactilia são mutações no gene *Shh*. Elas estão em um elemento regulatório de ação cis, longe da região codificante, que controla a expressão de *Shh* no desenvolvimento do broto dos membros. Os dedos extras são induzidos pela expressão de *Shh* em uma parte do membro onde o gene não é normalmente expresso. As mutações nos elementos regulatórios de ação cis têm duas propriedades importantes que são distintas das mutações nas regiões codificantes. Primeiro, como elas afetam a regulação em cis, os fenótipos em geral são dominantes. Segundo, como apenas um dos vários elementos regulatórios de ação cis pode ser afetado, outras funções podem ser completamente normais. A polidactilia pode ocorrer sem nenhum dos problemas desenvolvimentais colaterais. As mutações codificantes em *Shh*, entretanto, dizem uma história diferente, como veremos na próxima seção.

Holoprosencefalia

As mutações na região codificante *Shh* também foram identificadas. As alterações conseqüentes na proteína *Shh* estão associadas a uma síndrome chamada de *holoprosencefalia*, na qual as anormalidades ocorrem no tamanho do cérebro, na formação do nariz e em outras estruturas da linha média. Essas anomalias parecem ser contrapartes menos graves dos defeitos desenvolvimentais observados em camundongos mutantes homozigotos *Shh*. De fato, as crianças afetadas vistas em clínicas são heterozigotas. Uma cópia de um gene normal *Shh* parece ser insuficiente para o desenvolvimento da linha média normal (o gene é *haploinsuficiente*). Os fetos

humanos homozigotos para a perda de função *Shh* muito provavelmente morrem na gestação com mais defeitos graves.

A holoprosencefalia não é causada exclusivamente por mutações *Shh*. *Shh* é um ligando em uma via de transdução de sinal. Como pode ser esperado, as mutações nos genes que codificam outros componentes da via afetam a eficiência de sinalização de *Shh*, e estão associadas também à holoprosencefalia. Vários componentes da via *Shh* humana foram primeiro identificados como homólogos de membros da via da mosca, demonstrando novamente a conservação das ferramentas e o poder dos sistemas-modelo para as descobertas biomédicas.

Câncer como uma doença do desenvolvimento



Nos animais de vida longa, como nós mesmos e outros mamíferos, o desenvolvimento não termina ao nascimento ou ao final da adolescência. Os tecidos e vários tipos celulares estão constantemente sendo renovados. A manutenção de várias funções orgânicas depende do crescimento controlado e diferenciação de células que substituem as que foram descartadas ou, de outro modo, morreram. A manutenção de tecidos e órgãos geralmente é controlada por vias de sinalização. As mutações herdadas ou espontâneas nos genes que codificam os componentes dessas vias podem perturbar a organização tissular e contribuir para a perda de controle da proliferação celular. Como a proliferação celular descontrolada é uma característica do câncer, a formação de cânceres pode ser uma conseqüência. O câncer, portanto, é uma doença do desenvolvimento, um produto do processo normal de desenvolvimento que deu errado.

Alguns dos genes associados a tipos de cânceres humanos são membros compartilhados das ferramentas animais. Por exemplo, o gene *patched* codifica um receptor para as proteínas de sinalização Hedgehog. Além de causarem distúrbios desenvolvimentais herdados, tais como polidactilia e holoprosencefalia, as mutações no gene humano *patched* estão associadas à formação de uma variedade de cânceres. Cerca de 30 a 40% dos pacientes com um distúrbio genético dominante chamado de *síndrome do nevus de célula basal* (BCNS) levam mutações *patched*. Essas pessoas são fortemente predispostas a desenvolver um câncer de pele conhecido como carcinoma de célula basal. Eles também têm uma incidência muito aumentada de meduloblastoma, uma forma mortal de tumor cerebral. Hoje, cada vez mais cânceres estão associados a perturbações de vias de transdução de sinal, vias que foram

Quadro 12.2 Alguns Genes Ferramentais Têm Papéis no Câncer

	Gene de mosca	Gene de mamífero	Tipo de câncer
Componentes de via de sinalização			
Sem asas	<i>armadillo</i>	β - <i>catenin</i>	Cólon e pele
	<i>D.TCF</i>	<i>TLF</i>	Cólon
Hedgehog	<i>cubitus interruptus</i>	<i>Gli1</i>	Carcinoma de célula basal
	<i>patched</i>	<i>patched</i>	Carcinoma de célula basal, meduloblastoma
	<i>smoothened</i>	<i>smoothened</i>	Carcinoma de célula basal
Notch	<i>Notch</i>	<i>hNotch1</i>	Leucemia, linfoma
Receptor EGF	<i>torpedo</i>	<i>C-erbB-2</i>	Mama e cólon
Decapentaplégico/TFG- β	<i>Medea</i>	<i>DPC4</i>	Pancreático e cólon
Toll	<i>dorsal</i>	<i>NF-κB</i>	Linfoma
Outro	<i>extradenticle</i>	<i>Pbx1</i>	Leucemia aguda de célula pré-B

primeiro elucidadas por triagens genéticas sistemáticas de padrões mutantes nas moscas-das-frutas (Quadro 12.2). As descobertas de ligações entre mutações dos genes de vias de transdução de sinal e cânceres humanos facilitaram muito o estudo da biologia do câncer e o desenvolvimento de novas terapias. Por exemplo, cerca de 30% dos camundongos heterozigotos para uma mutação direcionada no gene *patched* desenvolvem meduloblastoma. Esses camundongos servem, portanto, como um excelente modelo para a biologia das doenças humanas e uma plataforma de testes para terapia. Muitos dos novos agentes anticâncer empregados hoje em dia são de

fato direcionados para componentes das vias de transdução de sinal que são perturbadas em alguns tipos de tumores. É justo dizer que mesmo os mais otimistas e imaginativos pesquisadores não esperavam que a descoberta das ferramentas genéticas para construir uma mosca tivessem efeitos tão amplos na compreensão do desenvolvimento humano e da doença. Mas tais dividendos não-previstos são familiares na história recente da pesquisa genética básica. O advento da medicina geneticamente construída, os anticorpos monoclonais para diagnóstico e terapia, e os testes de medicina forense tiveram, todos, origens similares em investigações aparentemente não-relacionadas.

Resumo

No Capítulo 10, mencionamos a frase de Jacques Monod de que “o que é verdade para *E. coli* também é verdade para o elefante”. Agora que vimos os processos regulatórios que fazem os vermes, moscas, camundongos e elefantes, diríamos que ele estava certo? Se Monod estivesse se referindo ao princípio de que a transcrição gênica é controlada por proteínas reguladoras específicas de seqüências, vimos que o repressor bacteriano Lac e as proteínas Hox de moscas de fato agem similarmente. Além disso, suas proteínas de ligação ao DNA têm o mesmo tipo de motivo. Os *insights* fundamentais que Jacob e Monod tiveram quanto ao papel central do controle da transcrição gênica na fisiologia bacteriana e o que eles esperavam que seria aplicado à diferenciação celular e desenvolvimento em organismos multicelulares complexos, surgiram em muitos aspectos no controle genético do desenvolvimento animal. Muitas características em eucariontes unicelulares e multicelulares, entretanto, não são encontradas em bactérias e seus vírus. Os geneticistas e biólogos moleculares descobriram as funções dos íntrons, a recomposição do RNA, vários e distantes elementos reguladores de ação cis, cromatina, recomposição alternativa e, mais recentemente, miRNA. Ain-

da, central ao controle genético do desenvolvimento está o controle da expressão gênica diferencial. Este capítulo apresentou uma visão geral da lógica e os mecanismos para o controle da expressão gênica e desenvolvimento em algumas espécies-modelo. Nós nos concentramos nos genes ferramentais para os processos de desenvolvimento animal e nos mecanismos que controlam a organização de características importantes do plano corpóreo, o estabelecimento dos eixos corpóreos, a segmentação e a identidade segmentar. Embora exploremos apenas um número modesto de mecanismos regulatórios em profundidade, e apenas algumas espécies, as similaridades na lógica regulatória e mecanismos nos permitem identificar alguns temas gerais quanto ao controle genético do desenvolvimento.

1. A despeito das grandes diferenças em aspecto e anatomia, os animais têm em comum genes ferramentais que controlam o desenvolvimento. Essas ferramentas são uma pequena fração de todos os genes no genoma, e a maioria desses genes ferramentais controlam fatores de transcrição e componentes de vias de transdução de sinal. Genes ferramentais individuais

tipicamente têm várias funções e afetam o desenvolvimento de estruturas diferentes em estágios diferentes.

2. *O desenvolvimento do embrião em desenvolvimento e de suas partes corpóreas ocorre em uma progressão espacial e temporalmente ordenada.* Os domínios dentro do embrião são estabelecidos pela expressão dos genes ferramentais que marcam progressivamente subdivisões mais finas ao longo de ambos os eixos embrionários.
3. *Padrões espacialmente restritos de expressão gênica são produtos de regulação combinatória.* Cada padrão de expressão gênica tem uma base causal precedente. Novos padrões são gerados por estímulos combinados de padrões precedentes. Nos exemplos apresentados neste capítulo, o posicionamento das faixas da regra dos pares e a restrição da expressão de apêndices de genes reguladores a genes individuais requer a integração de numerosos estímulos reguladores positivos e negativos por elementos regulatórios de ação cis.

A regulação pós-transcricional no RNA adiciona outra camada de especificidade ao controle da expressão gênica. A recomposição alternativa de RNA e o controle traducional por proteínas e miRNA também contribuem para o controle espacial e temporal da expressão de genes ferramentais.

O controle combinatório é a chave para a *especificidade* e *diversidade* da expressão gênica e funcionamento de genes ferramentais. Com relação à especificidade, mecanismos combinatórios fornecem os meios para localizar a

expressão gênica em discretas populações de células usando estímulos que não são específicos ao tipo de célula ou tecido. A ação de proteínas ferramentais pode assim ser bastante específica em diferentes contextos. Com relação à diversidade, os mecanismos combinatórios fornecem os meios para gerar uma variedade ilimitada de padrões de expressão gênica.

4. *A modularidade dos elementos regulatórios de ação cis permite o controle espacial e temporal independentemente da expressão e do funcionamento de genes ferramentais.* Do mesmo modo com que os operadores e UAS dos procariontes e eucariontes simples atuam como interruptores no controle fisiológico da expressão gênica, os elementos regulatórios de ação cis dos genes ferramentais atuam como interruptores no controle desenvolvimental da expressão gênica. A característica distintiva dos genes ferramentais é a presença típica de vários elementos reguladores de ação cis que controlam a expressão gênica em diferentes domínios espaciais e em diferentes estágios do desenvolvimento. A regulação espacial e temporal independente da expressão gênica permite que os genes ferramentais tenham funções diferentes, mas específicas, em contextos diferentes. Nesse contexto, não é adequado nem preciso descrever uma função de genes ferramentais apenas em relação à proteína (ou miRNA) que eles codificam, pois a função do produto gênico quase sempre depende do contexto no qual ele é expresso.

► Termos-chave

complexo gênico	gene de polaridade segmentar	homeobox
estrutura serialmente reiterada	gene de regra dos pares	homeodomínio
gene de efeito materno	gene gap	informação posicional
gene de manutenção	gene <i>Hox</i>	zigoto

► Problemas Resolvidos

Problema resolvido 1. O gene *Bicoid* (*bcd*) é um gene de efeito materno necessário para o desenvolvimento da região anterior de *Drosophila*. Uma mãe heterozigota para a deleção *bcd* tem só uma cópia do gene *bcd*. Com o uso de elementos *P* para inserir cópias do gene clonado *bcd*⁺ no genoma por transformação, é possível produzir mães com cópias extras do gene. Os embriões iniciais de *Drosophila* desenvolvem uma indentação chamada de sulco cefálico que é mais ou menos perpendicular ao eixo corpóreo (A-P) ântero-posterior longitudinal. Na prole de mães apenas com uma cópia de *bcd*⁺, esse sulco é muito próximo da ponta anterior, ficando em uma posição de um sexto da distância entre a ponta anterior e a posterior. Na prole de diplóides tipo selvagem padrão (tendo duas cópias de *bcd*⁺, o sulco cefálico surge mais posteriormente, em uma posição a um quinto da distância da ponta anterior para a posterior do embrião.

Na prole de mães com três cópias de *bcd*⁺, é ainda mais posterior. À medida que são acrescentadas doses gênicas adicionais, o sulco cefálico move-se mais e mais posteriormente, até que, na prole de mães com seis cópias de *bcd*⁺, está a uma distância média ao longo do eixo A-P do embrião. Explique o efeito da dosagem gênica de *bcd*⁺ na formação do sulco cefálico à luz da contribuição que *bcd* faz para uma formação de padrão A-P.

SOLUÇÃO

A determinação de partes ântero-posteriores do embrião é controlada por um gradiente de concentração da proteína Bicoid. O sulco desenvolve-se em uma concentração crítica de *bcd*. À medida que a dosagem gênica de *bcd*⁺ (e, portanto, a concentração da proteína Bicoid) diminui, o sulco move-se anteriormente; à medida que a dosagem gênica aumenta, o sulco muda posteriormente.

Problemas

PROBLEMAS BÁSICOS

1. *Gooseberry*, *runt*, *knirps* e *Antennapedia*. Para um geneticista de *Drosophila*, o que são? Como diferem?
2. Descreva o padrão de expressão do gene *eve* de *Drosophila* no embrião inicial.
3. Contraste a função dos genes homeóticos com a dos genes da regra dos pares.
4. Quando um embrião é homozigoto mutante para o gene gap *Kr*, a quarta e quinta faixas do gene da regra dos pares *ftz* (contando da ponta anterior) não se formam normalmente. Quando o gene gap *kni* é mutante, a quinta e sexta faixas *ftz* não se formam normalmente. Explique esses resultados com relação a como o número de segmentos é estabelecido no embrião.
5. Alguns dos genes *Hox* de mamíferos foram mostrados sendo mais similares a um dos genes *Hox* de insetos do que outros. Descreva um enfoque experimental que lhe permita demonstrar esse achado em um contexto funcional.
6. As três proteínas de homeodomínio ABD-B, ABD-A e UBX são codificadas por genes dentro do complexo *Bithorax* de *Drosophila*. Nos embriões tipo selvagem, o gene *Abd-B* é expresso nos segmentos abdominais posteriores, *Abd-A* no meio dos segmentos abdominais e *Ubx* nos segmentos torácicos abdominais anterior e posterior. Quando o gene *Abd-B* é deletado, *Abd-A* é expresso tanto no segmento abdominal médio quanto no posterior. Quando *Abd-A* é deletado, *Ubx* é expresso no tórax posterior e nos segmentos abdominais médio e anterior. Quando *Ubx* é deletado, os padrões de *Abd-A* e *Abd-B* não são mudados do tipo selvagem. Quando ambos, *Abd-A* e *Abd-B*, são deletados, *Ubx* é expresso em todos os segmentos do tórax posterior à extremidade posterior do embrião. Explique essas observações, levando em consideração o fato de que os genes gap controlam os padrões iniciais de expressão dos genes homeóticos.
7. Como você pode dizer se um gene é necessário zigoticamente e se ele tem efeito materno?
8. Considerando a formação dos eixos A-P e D-V em *Drosophila*, notamos que, para mutações tais como *bcd*, as mães mutantes homozigotas produzem uniformemente proles mutantes com defeitos de segmentação. Esse resultado é sempre verdadeiro, seja a prole *bcd⁺/bcd* ou *bcd/bcd*. Algumas outras mutações letais de efeito materno são diferentes, pois o fenótipo mutante pode ser “recuperado” introduzindo um alelo tipo selvagem do gene do pai. Em outras palavras, para tais letais recuperáveis de efeito materno, os animais *mut⁺/mut* são normais, enquanto os animais *mut/mut* têm o defeito mutante. Explique a diferença entre mutações letais de efeito materno recuperáveis e não-recuperáveis.
9. Suponha que você isola uma mutação afetando a padronização A-P do embrião de *Drosophila* na qual cada outro segmento da larva mutante em desenvolvimento está faltando.
 - a. Você consideraria essa mutação como sendo uma mutação no gene gap, um gene de regra dos pares, um

gene de segmento de polaridade ou um gene de identidade de segmento?

- b. Você clonou um pedaço de DNA que contém quatro genes. Você pode usar a expressão paterna espacial de seu mRNA em um embrião tipo selvagem para identificar o que representa um gene candidato para a mutação descrita?
 - c. Suponha que você identificou o gene candidato. Se você agora examina o padrão de expressão espacial de seu mRNA em um embrião que é homozigoto mutante para o gene gap *Krüppel*, você esperaria ver um padrão de expressão normal? Explique.
10. Como se forma o gradiente da proteína Bicoid?
 11. Em um embrião de uma fêmea mutante homozigota *Bicoid*, que classe(s) de expressão é(são) anormal(is)?
 - a. Genes gap
 - b. Genes da regra dos pares
 - c. Genes de polaridade segmentar
 - d. genes *Hox*

PROBLEMAS DESAFIADORES

12.
 - a. O gene *eyeless* é necessário para a formação do olho em *Drosophila*. Ele codifica um homeodomínio. O que você poderia prever sobre a função bioquímica da proteína Eyeless?
 - b. Onde você poderia prever que o gene *eyeless* seria expresso no desenvolvimento? Como você testaria sua previsão?
 - c. Os genes *Small eye* e *Aniridia* de camundongos e humanos, respectivamente, codificam proteínas com a forte similaridade de sequência à proteína Eyeless da mosca, e eles são denominados por seus efeitos no desenvolvimento do olho. Crie um teste para examinar se os genes de camundongos e humanos são funcionalmente equivalentes ao gene *eyeless* de mosca.
13. O gene *X* é expresso no cérebro em desenvolvimento, coração e pulmões de camundongo. As mutações que afetam seletivamente a função do gene *X* nesses três tecidos estão mapeadas em três regiões distintas (A, B e C, respectivamente) a 5' da região codificante de *X*.
 - a. Explique a natureza dessas mutações.
 - b. Desenhe um mapa do locus *X* coerente com a presente informação.
 - c. Como você testaria a função das regiões A, B e C?
14. Por que as mutações reguladoras do gene *Sonic hedgehog* são dominantes e viáveis? Por que as mutações codificantes causam mais defeitos generalizados?
15. Uma mutação ocorre no gene *doublesex* de *Drosophila*, impedindo que Tra se ligue ao RNA transcrito *dsx*. Quais seriam as consequências dessa mutação para a expressão da proteína Dsx em machos? Em fêmeas?
16. Você isola a mutação *glp-1* de *C. elegans* e descobre que a região do DNA que codifica o controle espacial (SCR) foi deletada. Qual será o padrão de expressão da proteína GLP-1 em um embrião de quatro células nos heterozigotos mutantes? Em homozigotos mutantes?

Genomas e Genômica

13



O genoma nuclear humano visto como um conjunto de DNA marcado. O DNA de cada cromossomo foi marcado com um corante que emite fluorescência de um comprimento de onda específico (produzindo uma cor específica). [Evelin Schrock and Thomas Ried.]

Perguntas Fundamentais

- Como as seqüências de mapa dos genomas são produzidas?
- Como a informação no genoma é decifrada?
- O que a Genômica comparativa pode revelar sobre a estrutura do genoma e a evolução?
- Como a disponibilidade da seqüência genômica afeta a análise genética?

Tópicos

- 13.1 A revolução genômica
- 13.2 Criando o mapa de seqüência de um genoma
- 13.3 Bioinformática: significado da seqüência genômica
- 13.4 A estrutura do genoma humano
- 13.5 Genômica comparativa
- 13.6 Genômica funcional e genética reversa

Em 1997, uma equipe de pesquisa na University of Munich, liderada por Svante Pääbo, relatou o seqüenciamento de uma região de 379 pb de DNA mitocondrial do osso da coxa do fóssil original de um neandertal, descoberto em 1856 (Figura 13.1). Esse seqüenciamento foi uma incrível conquista técnica. As moléculas de DNA quebram e acumulam modificações químicas com o passar do tempo, e, assim, apenas uma série de seqüências muito curtas podem ser decifradas e ligadas. A quantidade de DNA mitocondrial presente na amostra era muito pequena, e a quantidade de DNA nuclear era desprezível. Além disso, os cientistas tinham que tomar muito cuidado para terem certeza de que a seqüência que eles obtiveram não era um contaminante de outros humanos modernos ou alguma outra fonte. Mais interessante é que a seqüência do fragmento de DNA mitocondrial indicou que os neandertais tornaram-se extintos sem contribuir com o DNA mitocondrial para os humanos modernos.

Menos de 10 anos depois, Pääbo e sua equipe, hoje no Max Planck Institute for Evolutionary Anthropology, em Leipzig, anunciaram que tinham obtido

Amostra de osso de neandertal para seqüenciamento de DNA

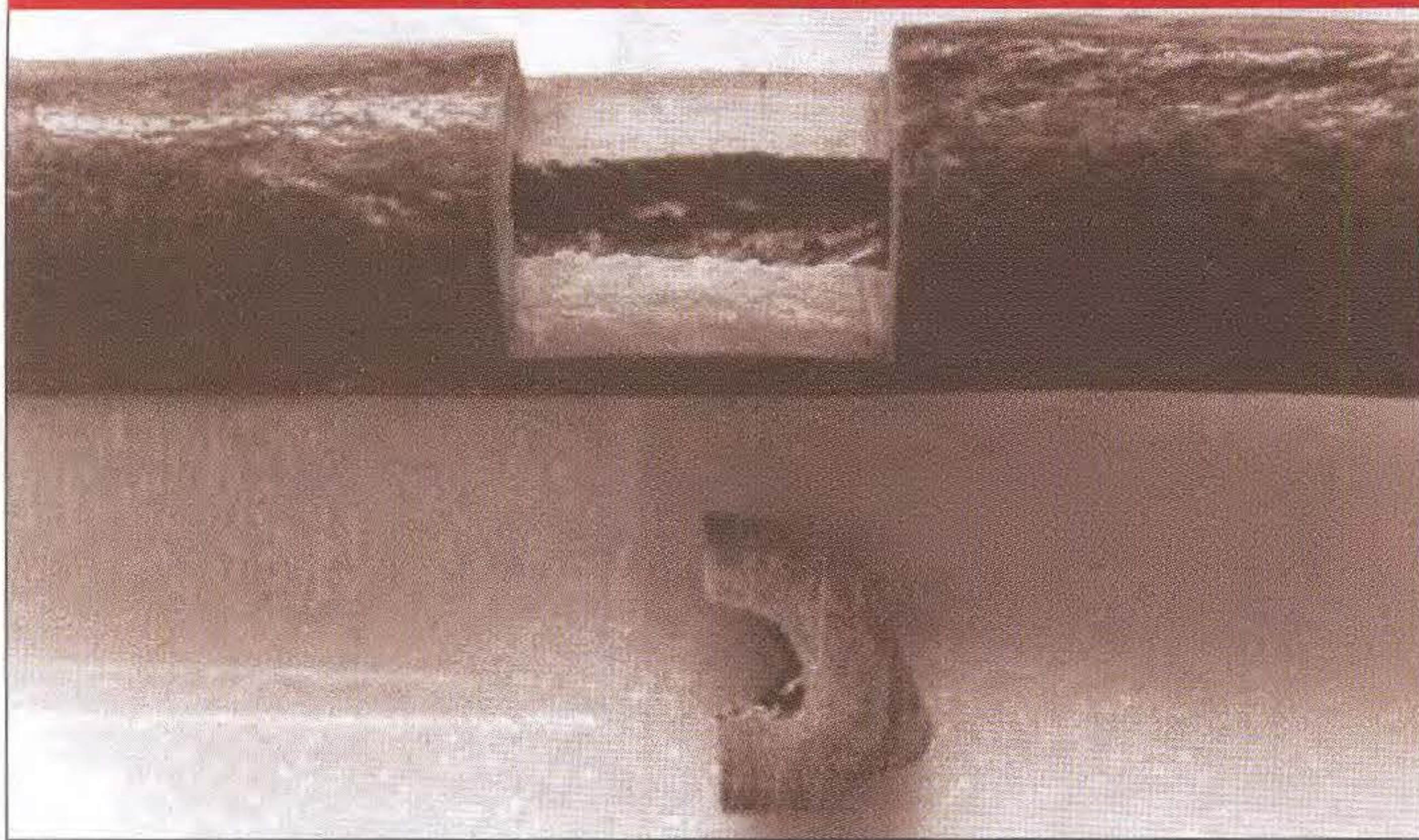


FIG. 13.1 Uma amostra (embaixo) foi removida do osso da coxa direita (em cima) de uma amostra original de neandertal para uso no seqüenciamento de DNA. [De M. Krings et al., "Neanderthal DNA Sequences and the Origin of Modern Humans," *Cell* 90, 1997, 19–30, Fig. 1.]

mais de 1 milhão de pares de bases de uma seqüência de DNA nuclear de uma amostra de neandertal. Além disso, eles estavam tentando obter seqüências genômicas completas não só de nosso primo recentemente extinto, mas também de 10 indivíduos neandertais.

Esses avanços na genética de neandertais ilustram os marcantes avanços na tecnologia e ambições da **Genômica**, o estudo de genomas em sua totalidade. O que começou como um filete de água hoje é uma inundação de dados. Em 1995, o genoma de 1,8 Mb (1,8 megabase) da bactéria *Haemophilus influenzae* foi o primeiro genoma de um organismo vivo a ser seqüenciado. Em 1996, veio o genoma de 12 Mb de *Saccharomyces cerevisiae*; em 1998, o genoma de 100 Mb de *C. elegans*; em 2000, o genoma de 180 Mb de *Drosophila melanogaster*; em 2001, o primeiro rascunho das 3.000 Mb do genoma humano; e, em 2005, o primeiro rascunho de nosso parente vivo mais próximo, o chimpanzé. Essas espécies são apenas uma pequena amostra; hoje temos as seqüências de mais de 300 genomas bacterianos, 50 genomas de fungos, várias plantas (*Arabidopsis*, arroz, por exemplo) e uma lista rapidamente crescente de mamíferos (rato, cão, gambá) e outros animais.

A Genômica revolucionou como a análise genética é feita e abriu caminhos para investigações que não eram concebíveis até alguns anos atrás. A maioria das análises genéticas que consideramos até agora empregam um enfoque direto para análise genética e processos biológicos. Isto é, a análise começa primeiro triando mutantes que afetam algum fenótipo observável, e a caracterização desses mutantes acaba levando ao nível do gene e à função do DNA, RNA e seqüências proteicas. Em contraste, saber as seqüências inteiras de DNA do genoma de um organismo permitiu que os geneticistas atuassem em ambas as direções: direto do fenótipo ao gene e, ao reverso, do gene ao fenótipo. Sem exceção, as seqüências genômicas revelam muitos genes que não eram detectados pela análise mutacional clássica. Usando a chamada genética

reversa, os geneticistas hoje podem estudar sistematicamente os papéis de tais genes antes não-identificáveis. Além disso, a falta do estudo genético clássico não é mais um impedimento para a investigação genética dos organismos. As fronteiras da análise experimental estão indo além do número modesto de organismos-modelo longamente explorados.

As análises de genomas inteiros hoje permeiam todos os cantos das pesquisas biológicas. Em Genética Humana, a Genômica está fornecendo novos modos para localizar genes que contribuem para muitas doenças genéticas determinadas por combinações complexas de fatores genéticos. Em sistemas-modelo, a disponibilidade de seqüências genômicas para espécies há muito estudadas e seus parentes acelerou muito a identificação gênica, a análise da função gênica e a caracterização de elementos não-codificantes do genoma. Novas tecnologias para a análise global genômica do papel fisiológico de todos os produtos gênicos está ativando o desenvolvimento do novo campo da *Biologia de Sistemas*. Por uma perspectiva evolucionária, a Genômica fornece uma visão detalhada de como os genomas divergiram e se adaptaram no tempo geológico. Nas pesquisas ecológicas, os biólogos estão desenvolvendo novos métodos para avaliar a distribuição de organismos com base na detecção da presença e concentração de genomas diferentes em amostras naturais. Na medicina humana, hoje podemos antever quando a seqüência genômica será uma parte padrão de seu registro médico.

A seqüência de DNA do genoma é o ponto de partida para todo um novo conjunto de análises destinado a compreender a estrutura, função e evolução do genoma e seus componentes. Neste capítulo, enfocaremos três aspectos principais da análise genômica:

- **Bioinformática**, que analisa o conteúdo de informações de genomas inteiros. Essa informação inclui os números e tipos de genes e produtos gênicos, bem como os sítios de ligação no DNA e RNA que permitem que os produtos funcionais sejam feitos na época e local corretos.
- **Genômica comparativa**, que considera os genomas de espécies próximas e distantemente correlatas para um insight evolutivo, e permite que seqüências conservadas sejam usadas como um guia para analisar o funcionamento gênico.
- **Genômica funcional**, que usa uma variedade crescente de métodos, incluindo a Genética reversa, para compreender a função gênica e delinear as redes de interações gênicas e de proteínas em processos biológicos.

► 13.1 A Revolução Genômica

Após o desenvolvimento da tecnologia do DNA recombinante nos anos de 1970, as pesquisas de laboratórios tipicamente fizeram a clonagem e o seqüenciamento de um gene de cada vez, e, então, apenas após ter encontrado primeiro algo interessante sobre esse gene a partir de uma análise mutacional clássica. As etapas no procedimento de ter um mapa genético clássico de um locus para isolar o DNA codificante de um gene (*clonagem*) e determinar sua seqüência geralmente eram numerosas e demoradas. Nos anos de

1980, alguns cientistas perceberam que uma grande equipe de pesquisadores em um esforço conjunto poderia clonar e sequenciar *todo* o genoma de um organismo selecionado. Tais **Projetos Genoma** tornariam então os clones e a sequência publicamente disponíveis. Um atrativo desse recurso é que, quando os pesquisadores se tornaram interessados em um gene de uma espécie cujo genoma foi sequenciado, eles só precisavam descobrir onde estava situado esse gene no mapa do genoma para serem capazes de focar sua sequência e potencialmente sua função. Por esse meio, um gene pode ser caracterizado muito mais rapidamente do que pela clonagem e sequenciamento a partir do zero, um projeto que poderia levar vários anos para ser feito. Esse enfoque mais rápido hoje é uma realidade para todos os organismos-modelo. De modo similar, em genética humana, a sequência genômica pode ajudar a identificar genes causadores de doenças.

Em uma perspectiva mais ampla, os projetos genômicos tinham o atrativo de dar um vislumbre dos princípios pelos quais os genomas são feitos. A obtenção de uma sequência genômica é como ter desenterrado algum escrito em uma linguagem que não podemos decifrar. O genoma humano, por exemplo, tem 24 fileiras de pares de bases, representando os cromossomos X e Y e os 22 autossomos. No total, o genoma humano contém 3 bilhões de pares de bases de DNA. Embora possamos nos convencer de que compreendemos um único gene de interesse, o grande desafio da Genômica hoje em dia é a literatura genômica: Como lemos a armazenagem de informações codificadas na sequência de tais genomas?

As técnicas básicas necessárias para sequenciar genomas inteiros já estavam disponíveis nos anos de 1980, incluindo plasmídeos bacterianos e cromossomos de bacteriófagos (usados como veículos para clonagem de DNA), a reação em cadeia da polimerase (PCR) para amplificar os genes, e as máquinas de sequenciamento de DNA. Mas a escala que foi necessária para sequenciar um genoma complexo foi, como projeto de engenharia, muito além da capacidade de pesquisa da comunidade. A Genômica, no final dos anos de 1980 e 1990, desenvolveu-se de grandes centros de pesquisa que podem integrar essas tecnologias elementares em uma linha de produção de nível industrial. Esses centros desenvolveram robótica e automação para os muitos milhares de etapas de clonagem e milhões de reações de sequenciamento necessários para montar a sequência de um organismo complexo. Com esses centros no lugar, o final dos anos de 1990 e 2000 foram a idade de ouro do sequenciamento do genoma. A taxa de sequenciamento do genoma continuou a acelerar-se. Novas tecnologias combinando microfluidos e fibras ópticas podem obter mais de 25 milhões de sequências de bases em um dia de trabalho em um único instrumento.

A Genômica, ajudada pelo crescimento explosivo da tecnologia da informação, estimulou os pesquisadores a desenvolver meios de experimentação, bem como a análise computacional, o genoma como um todo, em lugar de um gene de cada vez. Também foi demonstrado o valor de coletar dados em larga escala antes de seu uso, com grande potencial para abordar problemas específicos de pesquisa. A Genômica também mudou a sociologia da pesquisa biológica, demons-

trando o valor das grandes redes colaborativas de pesquisa como um complemento aos pequenos laboratórios de pesquisa independente (que ainda florescem). Esses efeitos só aumentarão à medida que surgirem mais informações, tecnologia e *insight*. Na última seção deste capítulo, exploraremos alguns meios que a Genômica, hoje, usa nas pesquisas genéticas básica e aplicada. Em capítulos subsequentes, veremos como a Genômica está catalisando avanços na compreensão da dinâmica da mutação, recombinação e evolução.

A caracterização de genomas inteiros é importante para uma compreensão fundamental dos princípios que funcionam em organismos vivos, e para a descoberta de novos genes, tais como os que têm papéis nas doenças humanas.

► 13.2 Criando o Mapa de Sequência de um Genoma

Quando as pessoas encontram um novo território, uma de suas primeiras atividades é criar um mapa. Essa prática foi verdadeira para exploradores, geógrafos, oceanógrafos e astrônomos, e é igualmente verdadeira para geneticistas. Os geneticistas usam muitos tipos de mapas para explorar o terreno de um genoma. Os exemplos são mapas de ligação baseados nos padrões de herança de genes alelos e mapas citogenéticos baseados na localização de características microscopicamente visíveis, tais como pontos de quebra de rearranjos.

O mapa de maior resolução é a sequência completa do DNA do genoma, isto é, a sequência completa de nucleotídeos A, T, C e G de cada dupla hélice no genoma. Como fazer um mapa completo da sequência de um genoma é um trabalho imenso, nunca visto antes na biologia, devem ser usadas novas estratégias, todas baseadas na automação.

Transformando leituras de sequências em um mapa de sequência

Você, provavelmente, já viu um ato de magia no qual o mágico corta uma página de jornal em muitos pedaços, mistura-os em sua cartola, diz algumas palavras mágicas e *voilà!* reaparece uma página inteira de jornal. Basicamente, é assim que são produzidos os mapas de sequência genômica. O enfoque é (1) quebrar o genoma em milhares a milhões de pequenos segmentos aleatórios, (2) ler a sequência de cada pequeno segmento, (3) encontrar computacionalmente a superposição entre os pequenos segmentos onde suas sequências são idênticas e (4) continuar a superposição de pedaços maiores até que todos os pequenos segmentos estejam ligados (Figura 13.2). Nesse ponto, um mapa de sequência de um genoma está montado.

Por que esse processo requer automação? Para compreender o porquê, vamos considerar o genoma humano, que contém cerca de 3×10^9 pb de DNA, ou 3 bilhões de pares de bases (3 pares de gigabases = 3 Gpb). Suponha que podemos purificar o DNA intacto de cada um dos 24 cromossomos

A lógica de criar um mapa de sequência do genoma

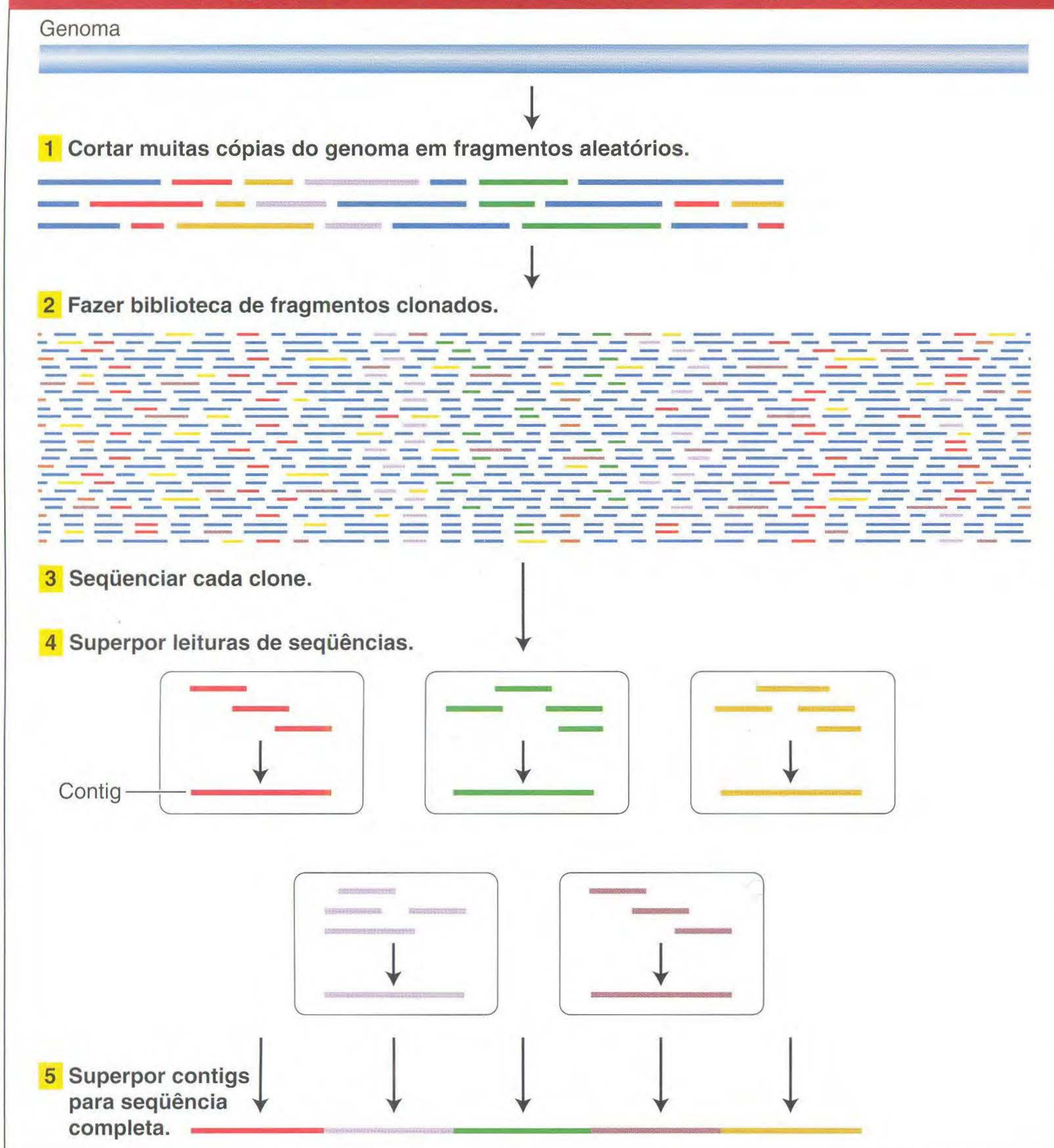


FIG. 13.2 Para criar um mapa de seqüência do genoma, várias cópias do genoma são cortadas em pequenos pedaços que são então clonados e seqüenciados. As leituras de seqüências resultantes são superpostas comparando seqüências idênticas em clones diferentes até que seja produzida uma seqüência de consenso de cada dupla hélice do DNA no genoma.

humanos (X, Y e os 22 autossomos), colocar separadamente cada uma dessas 24 amostras de DNA em uma máquina de seqüenciamento e ler diretamente suas seqüências de um telômero ao outro. Criar uma seqüência de mapa completa seria bem direto, como ler um livro com 24 capítulos, embora um livro muito longo com 3 bilhões de características (do tamanho de cerca de 3.000 romances). Infelizmente, tal máquina de seqüenciamento não existe. O seqüenciamento automatizado baseado em fluorescência do tipo discutido no Capítulo 20 é atualmente o mais perfeito na tecnologia de seqüenciamento de DNA. As reações individuais de seqüenciamento (chamadas de *leituras de seqüenciamento*) fornecem seqüências de letras que têm geralmente cerca de 600 bases de tamanho. Tais tamanhos são pequenos comparados com o DNA de um único cromossomo. Por exemplo, uma leitura individual tem apenas 0,0002% do maior cromossomo humano (cerca de 3×10^8 pb de DNA) e apenas cerca de 0,00002% de todo o genoma humano. Assim, o maior desafio enfrenta-

do pelo projeto genoma é a **montagem de seqüência**, isto é, juntar todas as leituras individuais em uma **seqüência de consenso**, uma seqüência para a qual há um consenso (ou concordância) que é uma autêntica representação da seqüência para cada uma das moléculas de DNA nesse genoma.

Vejamos esses números de um modo um pouco diferente para compreender a escala do problema. Como em qualquer observação experimental, as máquinas de seqüenciamento automatizado nem sempre dão leituras perfeitamente precisas da seqüência. A taxa de erro não é constante; ela depende de fatores tais como corantes que são ligados às moléculas seqüenciadas, pureza e homogeneidade da amostra inicial de DNA e seqüência específica de pares de bases na amostra de DNA. Assim, para garantir a precisão, os projetos genoma convencionalmente obtêm múltiplas (até 10) leituras independentes de seqüências de cada par de bases em um genoma. A cobertura de 10 vezes (representada por 10X) assegura que as chances de erro nas leituras não dão uma

falsa reconstrução da sequência de consenso. Considerando uma sequência média de leitura de cerca de 600 bases de DNA e um genoma humano de 3 bilhões de pares de bases, são necessários 50 milhões de leituras independentes para dar uma cobertura média de 10 vezes de cada par de bases. Entretanto, nem todas as leituras são bem-sucedidas; e, assim, o número de leituras tentadas é maior. Logo, a quantidade de informações e material a ser rastreado é enorme. Para tentar minimizar tanto o erro humano quanto a necessidade das pessoas em fazer tarefas altamente repetitivas, os laboratórios de projetos genoma implementaram a automação, computação com o uso de código de barras e sistemas de análise computadorizada sempre que praticáveis.

Por esses motivos, a preparação de clones, isolamento de DNA, eletroforese e protocolos de sequenciamento foram todos adaptados à automação. Por exemplo, uma conquista recente foi o desenvolvimento das máquinas de linha de produção de sequenciamento que funcionam noite e dia sem intervenção humana. Grandes genomas de mamíferos foram sequenciados por centros de genomas com muitas máquinas de sequenciamento funcionando em paralelo, produzindo até 150.000 leituras em um único dia. Um único centro de sequenciamento tem a capacidade de montar a sequência de um genoma de mamífero de 3 Gpb em 1 a 2 anos. A Figura 13.3 mostra uma linha de sequenciamento. Novos desenvolvimentos estão acelerando a taxa de sequenciamento e diminuindo os custos.

Quais são as metas de sequenciar um genoma? Primeiro, produzir uma sequência de consenso que seja uma representação verdadeira e precisa do genoma, começando com um organismo individual ou linhagem padrão da qual o DNA foi obtido. Essa sequência servirá então como sequência de

referência para a espécie. Agora sabemos que existem muitas diferenças na sequência de DNA entre indivíduos diferentes dentro de uma espécie, e mesmo entre os genomas contri- buídos pela mãe e pelo pai dentro de um único indivíduo diplóide. Assim, nenhuma sequência de genoma representa verdadeiramente o genoma de toda a espécie. Entretanto, a sequência do genoma serve como um padrão ou referência com a qual outras sequências podem ser comparadas, e ela pode ser analisada para determinar a informação codificada dentro do DNA, tal como o inventário dos RNA e polipeptí- deos codificados.

Como os manuscritos, as sequências genômicas podem variar de qualidade *rascunho* (o esquema geral está lá, mas existem erros tipográficos, erros gramaticais, espaços, seções que precisam de rearranjos, e assim por diante), até a quali- dade *final* (uma taxa muito baixa de erros tipográficos, algu- mas seções faltando, mas tudo que é atualmente possível foi feito para preencher essas seções), até a verdadeiramente *completa* (sem erros tipográficos, cada par de bases absolu- tamente correto de telômero a telômero). Nas seções seguin- tes, consideraremos os métodos gerais para a produção das montagens rascunho e sequência genômica terminada, bem como algumas das características dos genomas que desafiam os projetos de sequenciamento genômico.

Estabelecendo uma biblioteca genômica de clones

Existem duas estratégias gerais para montagem de sequência de um genoma. Uma é chamada de *sequenciamento shotgun de genoma inteiro* (WGS) e a outra é chamada de *sequencia- mento de clone ordenado*. Ambas as estratégias são baseadas na determinação da sequência de muitos segmentos genô-

O sequenciamento genômico hoje é automatizado

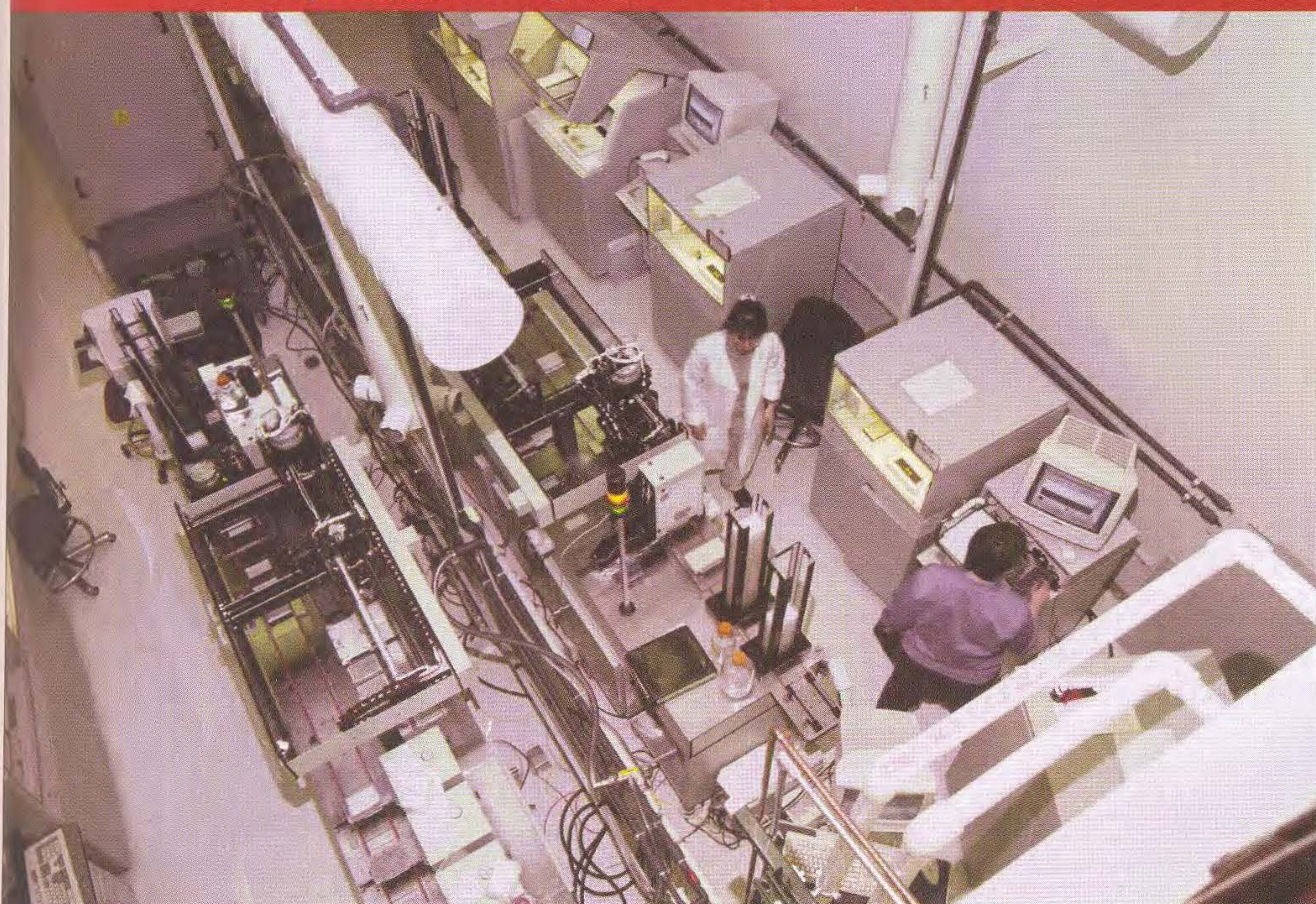


FIG. 13.3 Parte da linha de produção automatizada de um grande centro de sequen- ciamento do genoma huma- no. Todo esse equipamento é usado para o processamento rápido de um imenso número de clones para sequenciamento do DNA. [Copyright Bethany Ver- soy; todos os direitos reservados.]

micos de DNA que foram gerados pela quebra de DNA de longos cromossomos em muitos segmentos pequenos. Cada **enfoque** começa com a construção de **bibliotecas genômicas**, que são coleções desses curtos segmentos de DNA representando o genoma inteiro. Os curtos segmentos de DNA em tal biblioteca foram inseridos em um de vários tipos de cromossomos *acessórios* (elementos não-essenciais, tais como plasmídeos, vírus bacterianos modificados ou cromossomos artificiais) e propagados em micróbios, geralmente bactérias ou leveduras. Os cromossomos acessórios levando inserções de DNA são chamados de **vetores**.

Para gerar uma biblioteca genômica, um pesquisador primeiro usa *enzimas de restrição*, que clivam o DNA em seqüências específicas para retalhar o DNA genômico purificado. Algumas enzimas cortam o DNA em muitos lugares, enquanto outras cortam em menos lugares; assim, o pesquisador pode controlar se o DNA é cortado, em média, em pedaços maiores ou menores. Os fragmentos resultantes têm filamentos únicos pequenos de DNA em ambas as pontas. Esses fragmentos são então unidos ao DNA do cromossomo acessório, que também foi cortado com uma enzima de restrição e que tem pontas que são complementares às dos fragmentos genômicos. Para que todo o genoma seja representado, milhares a milhões de tais moléculas recombinantes diferentes devem ser geradas de múltiplas cópias do DNA genômico com cortes únicos aleatórios.

O conjunto resultante de moléculas recombinantes de DNA é então propagado, tipicamente por sua introdução em células bacterianas. Cada célula capta uma molécula recombinante. Então, cada molécula recombinante é replicada no crescimento normal e divisão de seu hospedeiro de modo que muitas moléculas idênticas do fragmento inserido são produzidas para uso ao analisar a seqüência de fragmentos de DNA. Como cada molécula recombinante é amplificada de uma célula individual, cada célula é um *clone* distinto. (Mais detalhes sobre clonagem de DNA são dados no Capítulo 20.)

A maioria das estratégias de seqüenciamento genômico são baseadas em clones. Com o uso de clones de uma biblioteca genômica, é obtida a seqüência de DNA genômico inserido imediatamente adjacente ao DNA vetor. Essa informação é então usada de modos diferentes para reunir a seqüência genômica, conforme tenha sido aplicado um enfoque *shotgun* ou ordenado.

Seqüenciamento de um genoma simples usando o enfoque *shotgun* de genoma inteiro

A lógica do seqüenciamento *shotgun* de genoma inteiro é: seqüenciar primeiro, mapear depois. Primeiro, as leituras das seqüências são obtidas de clones aleatoriamente selecionados de uma biblioteca de genoma inteiro sem nenhuma informação de onde esses clones mapeiam no genoma. Tal biblioteca é chamada de *biblioteca shotgun*. Então, essas leituras de seqüências são montadas em uma seqüência de consenso cobrindo todo o genoma combinando seqüências homólogas compartilhadas por leituras de clones superpostos.

O DNA bacteriano é essencialmente DNA de *cópia única*, sem seqüências repetidas. Portanto, qualquer determinada seqüência de DNA lida de um genoma bacteriano virá de um

único lugar nesse genoma. Além disso, um genoma bacteriano típico tem apenas alguns pares de megabases de DNA de tamanho. Devido a essas propriedades, o seqüenciamento WGS pode ser aplicado a genomas bacterianos de um modo direto.

Como são obtidas as seqüências? A reação de seqüenciamento começa de um primer de seqüência conhecida. Como a seqüência de um inserto clonado não é conhecida (e é a meta do exercício), os primers são baseados na seqüência do DNA vetor adjacente. Esses primers são usados para guiar a reação de seqüenciamento do inserto. Assim, regiões curtas em uma ou em ambas as pontas dos insertos genômicos podem ser seqüenciadas (Figura 13.4). Após o seqüenciamento, o produto é uma grande coleção de seqüências curtas aleatórias, algumas delas superpostas. As seqüências de leituras superpostas são reunidas em unidades chamadas de **seqüências contigs** (seqüências que são contíguas, ou que se tocam). Cada contig cobre uma grande região do genoma bacteriano. Com o uso do enfoque WGS, em julho de 2006, 320 espécies bacterianas foram totalmente seqüenciadas, e várias centenas de projetos adicionais de seqüenciamento bacteriano estão em progresso.

Usando o enfoque *shotgun* do genoma inteiro para criar uma seqüência-rascunho de um genoma complexo

Uma grande barreira em remontar uma seqüência de consenso de um genoma eucariótico é a existência de várias

Leituras finais de múltiplos insertos podem ser superpostas para produzir um contig

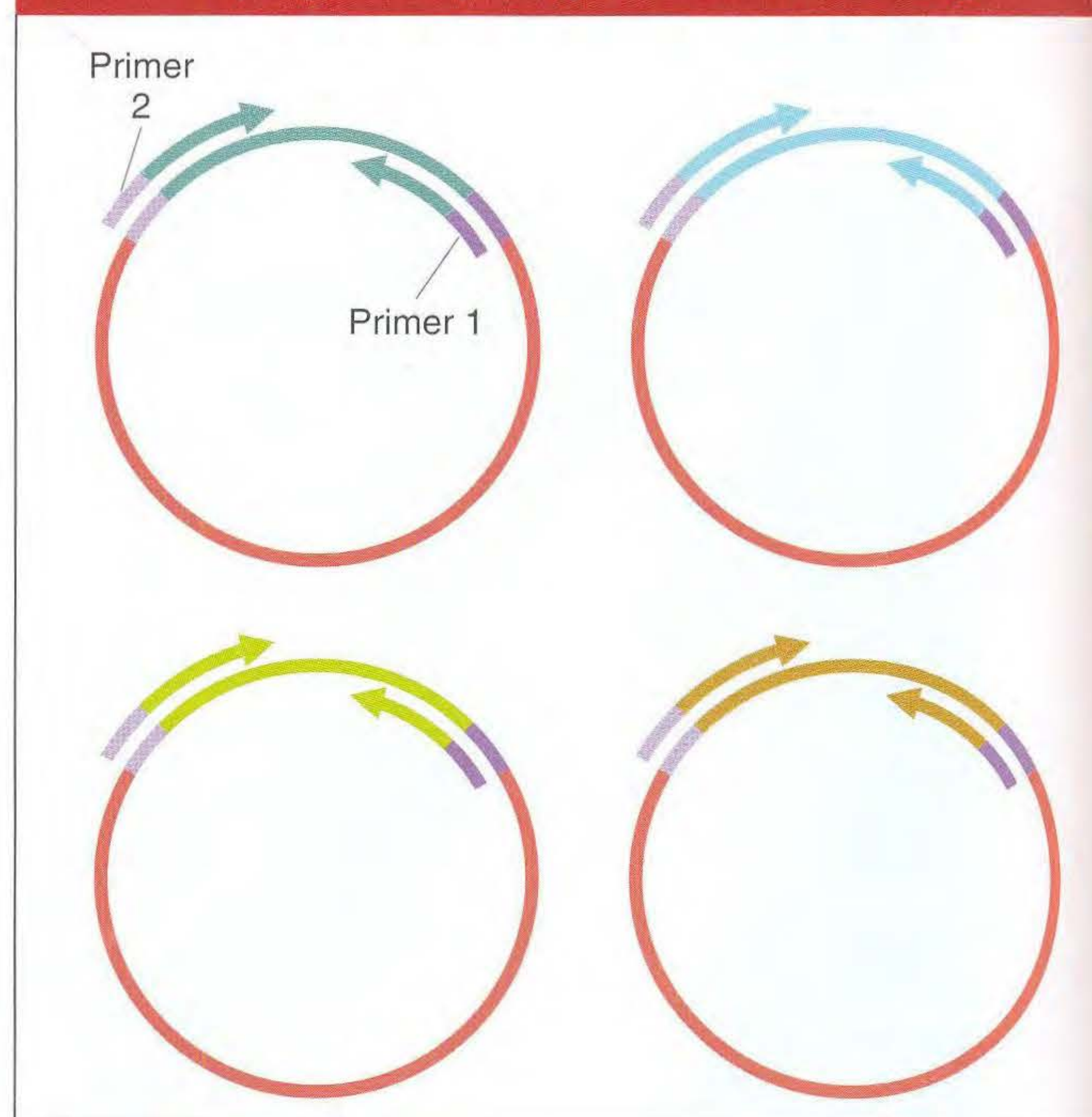


FIG. 13.4 Leituras de seqüenciamento são colhidas apenas das pontas de insertos clonados. (a) O uso de seqüências de dois sítios de iniciação diferentes, um em cada ponta do vetor, permite o seqüenciamento de até 600 pares de bases em cada ponta do inserto genômico. Se ambas as pontas do mesmo clone são seqüenciadas, as duas seqüências resultantes são chamadas de *leituras de pontas pareadas*.

classes de seqüências repetidas, algumas dispostas em tandem e outras dispersas. Por que elas são um problema para o seqüenciamento genômico? Em resumo, porque um seqüenciamento lido de DNA repetitivo ajusta-se a muitos locais no rascunho do genoma. Não raramente, uma seqüência repetida em tandem no total é maior que o tamanho de uma leitura máxima de seqüência. Nesse caso, não há modo de fazer uma ponte entre seqüências únicas adjacentes. Os elementos repetitivos dispersos podem causar alinhamento errado e atribuições de leituras de diferentes cromossomos ou partes diferentes do mesmo cromossomo.

O panorama de cromossomos eucarióticos inclui uma variedade de segmentos de DNA repetitivos. Esses segmentos são difíceis de alinhar como leituras de seqüências.

O seqüenciamento shotgun de genoma inteiro é particularmente bom para produzir seqüências com qualidade de rascunho de genomas complexos com muitas seqüências repetitivas. Como exemplo, consideraremos o genoma da mosca-das-frutas *D. melanogaster*, que foi inicialmente seqüenciado por esse método. O projeto começou com o seqüenciamento de bibliotecas de clones genômicos de tamanhos diferentes (2 kb, 10 kb, 150 kb). As leituras das seqüências foram obtidas de *ambas* as pontas dos insertos de clones genômicos e alinhadas por uma lógica idêntica à usada para o seqüenciamento WGS procariótico. Por essa lógica, as seqüências homólogas superpostas foram identificadas e os clones foram colocados em ordem, produzindo seqüências contigs, seqüências de consenso para esses trechos de cópia única do genoma. Entretanto, ao contrário da situação em bactérias, onde há apenas uma única cópia de DNA, os contigs correm para um segmento de DNA repetitivo que impede a montagem de contigs em um genoma inteiro. A seqüência de contigs tem um tamanho médio de cerca de 150 kb. O desafio então foi como grudar os milhares de tais seqüências de contigs em sua ordem e orientação corretas.

A solução para esse problema foi fazer uso de pares de leituras de seqüências de extremidades opostas dos insertos genômicos no mesmo clone. Essas leituras são chamadas de **leituras de**

Leituras de pontas pareadas podem ser usadas para unir duas seqüências contigs

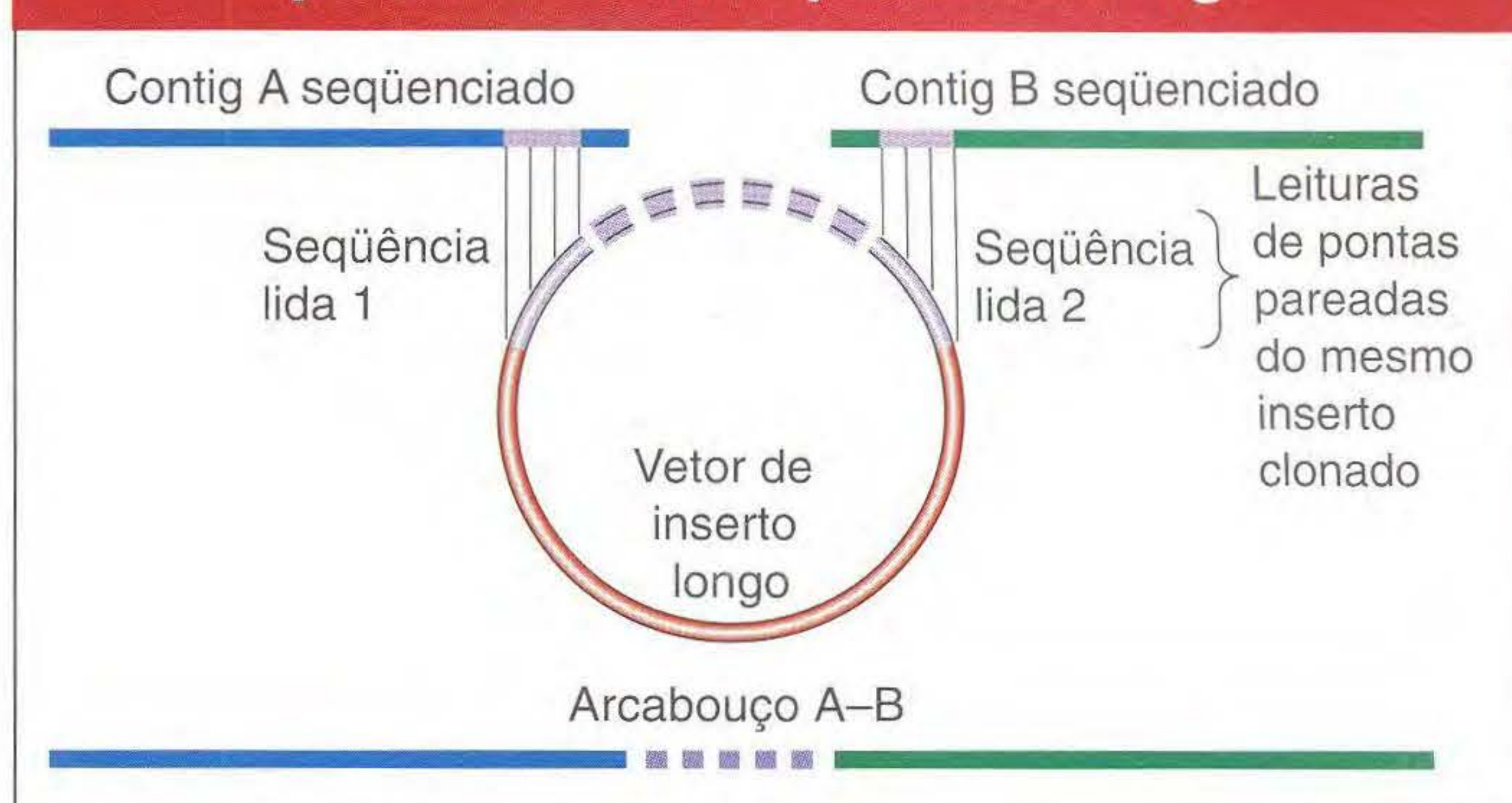


FIG. 13.5 Leituras de pontas pareadas podem ser usadas para unir duas seqüências contig em um único arcabouço ordenado e orientado.

pontas pareadas. A idéia foi encontrar leituras de extremidades pareadas que abrangiam espaços (gaps) entre duas seqüências contigs (Figura 13.5). Em outras palavras, se uma ponta de um inserto era parte de um contig e a outra ponta era parte de um segundo contig, então esse inserto deve envolver o espaço entre dois contigs, e os dois contigs estavam claramente perto um do outro. De fato, como o tamanho de cada clone era conhecido (isto é, veio de uma biblioteca contendo insertos genômicos de tamanho uniforme, seja uma biblioteca de 2 kb, 100 kb ou 150 kb), a distância entre as pontas de leituras era conhecida. Além disso, o alinhamento das seqüências dos dois contigs usando leituras de pontas pareadas determina automaticamente a orientação relativa de dois contigs. Desse modo, os contigs de uma cópia podiam ser unidos, embora com espaços onde residem os elementos repetitivos. Essas coleções espaçadas de seqüências contigs unidas são chamadas de **arcabouços** (às vezes também chamadas de **supercontigs**). Como a maioria das repetições de *Drosophila* são grandes (3–8 kb) e amplamente espaçadas (uma repetição em aproximadamente cada 150 kb), essa técnica foi extremamente efetiva em produzir uma seqüência rascunho corretamente montada da cópia única de DNA. Um resumo da lógica desse enfoque é mostrado na Figura 13.6.

Estratégia de montagem de seqüenciamento shotgun de genoma inteiro

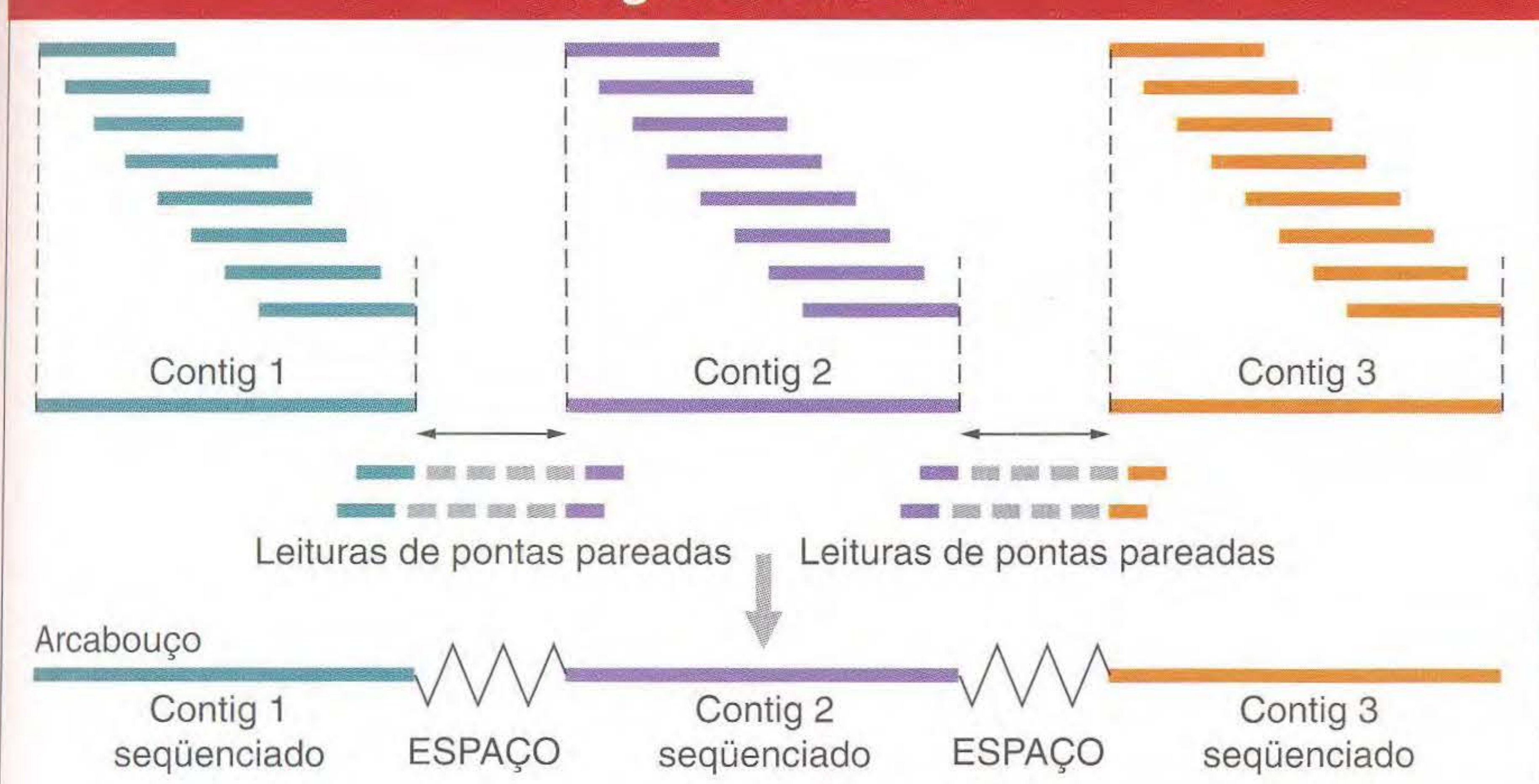


FIG. 13.6 Usando o enfoque shotgun de genoma inteiro, primeiro, a seqüência única se superpõe entre leituras de seqüências usadas para construir contigs. As leituras de pontas pareadas são então usadas para incluir espaços e ordenar e orientar contigs em unidades maiores, chamadas de *arcabouços*.

Usando o enfoque de clone ordenado para seqüenciar um genoma complexo

A lógica do seqüenciamento de clone ordenado é o oposto do enfoque shotgun do genoma inteiro: mapear primeiro, seqüenciar depois. Os insertos clonados individuais de uma biblioteca genômica são triados quanto a similaridades em,

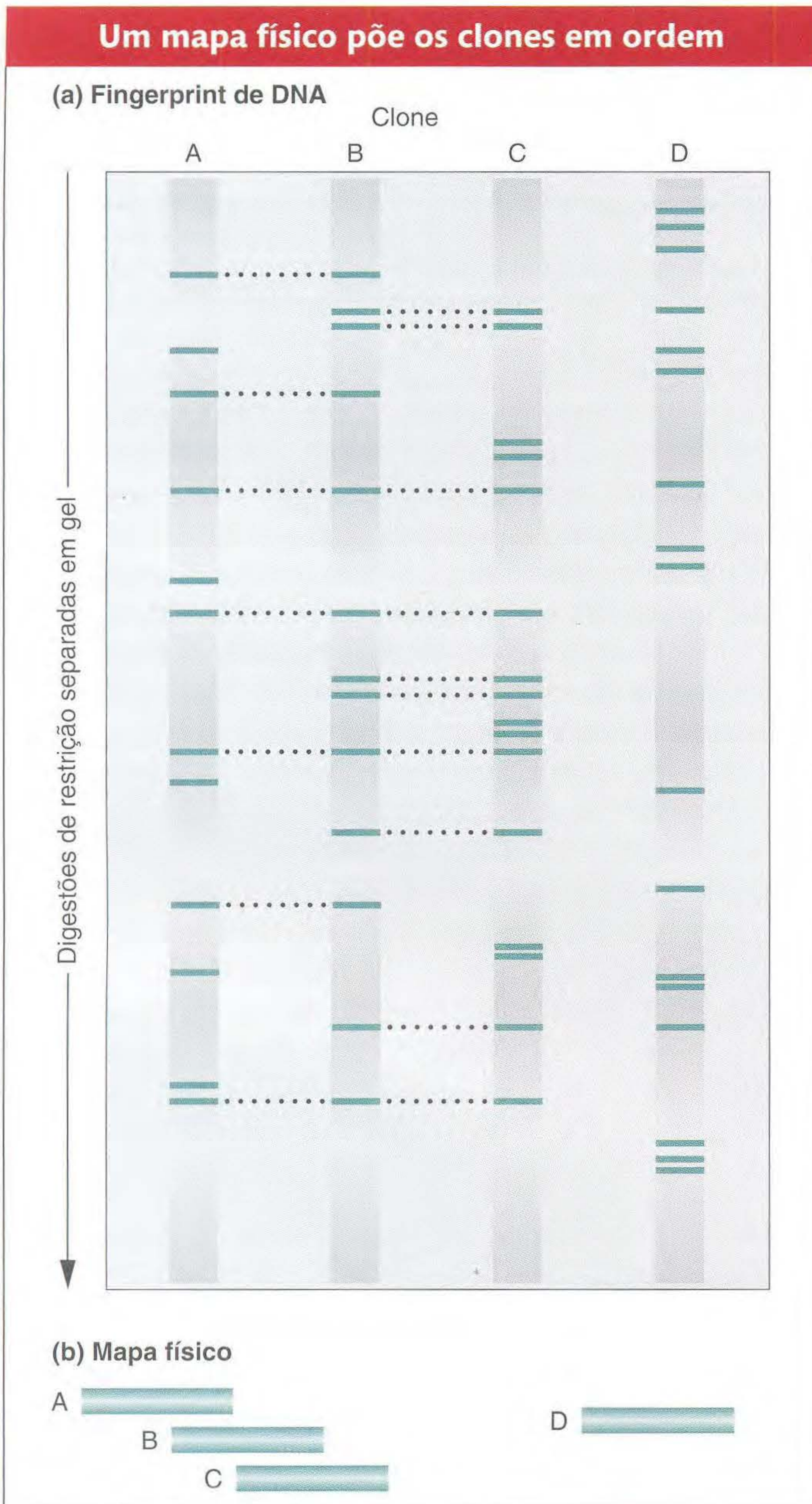


FIG. 13.7 Criando um mapa físico por mapeamento de fingerprint de clones. (a) Quatro clones foram digeridos com várias enzimas de restrição, e a mistura complexa resultante de fragmentos de restrição é separada com base no tamanho por eletroforese em gel. As bandas contendo os fragmentos são coradas para mostrar seu local. O número de bandas de tamanho idêntico para cada par de digestões é determinado. As digestões A e B compartilham mais de 50% das bandas, como as digestões B e C indicando que vieram de regiões superpostas do genoma. Várias bandas são encontradas em todas as três, A, B e C, sugerindo que alguma parte dos três clones se superpõe. (b) O mapa físico derivado dos dados na parte (a). O clone D é de outra parte do genoma, pois não se superpõe a nenhum dos outros três clones.

por exemplo, sítios de reconhecimento de enzimas de restrição, indicando que dois insertos se superpõem para formar um **clone contig** (Figura 13.7). A superposição nos diz que os insertos são contínuos no genoma. Esse procedimento resulta em um conjunto de clones ordenados e orientados que, juntos, cobrem o genoma. Tal conjunto ordenado e orientado de clones cobrindo todo o genoma é chamado de **mapa físico** do genoma. Aqui, a palavra “físico” é usada no sentido de que o mapa é composto de objetos reais (segmentos de DNA) que podem ser isolados e estudados em um tubo de ensaio.

Nas primeiras fases de um projeto genoma, os clones contigs representando segmentos separados no genoma são numerosos. Mas, à medida que mais clones são caracterizados, são encontrados clones que superpõem dois clones contíguos antes separados, e esses *clones unidos* então permitem a fusão dos dois clones contigs em um grande contig. Esse processo de junção de contigs continua até que o conjunto de clones contig seja igual ao número de cromossomos. Nesse ponto, se cada clone contig estende-se para fora dos telômeros de seu cromossomo, o mapa físico está completo.

O mapeamento físico continua ordenando os clones em grupos superpostos chamados de clones contigs. À medida que mais dados se acumulam, esses clones contigs se estendem ao longo de todos os cromossomos.

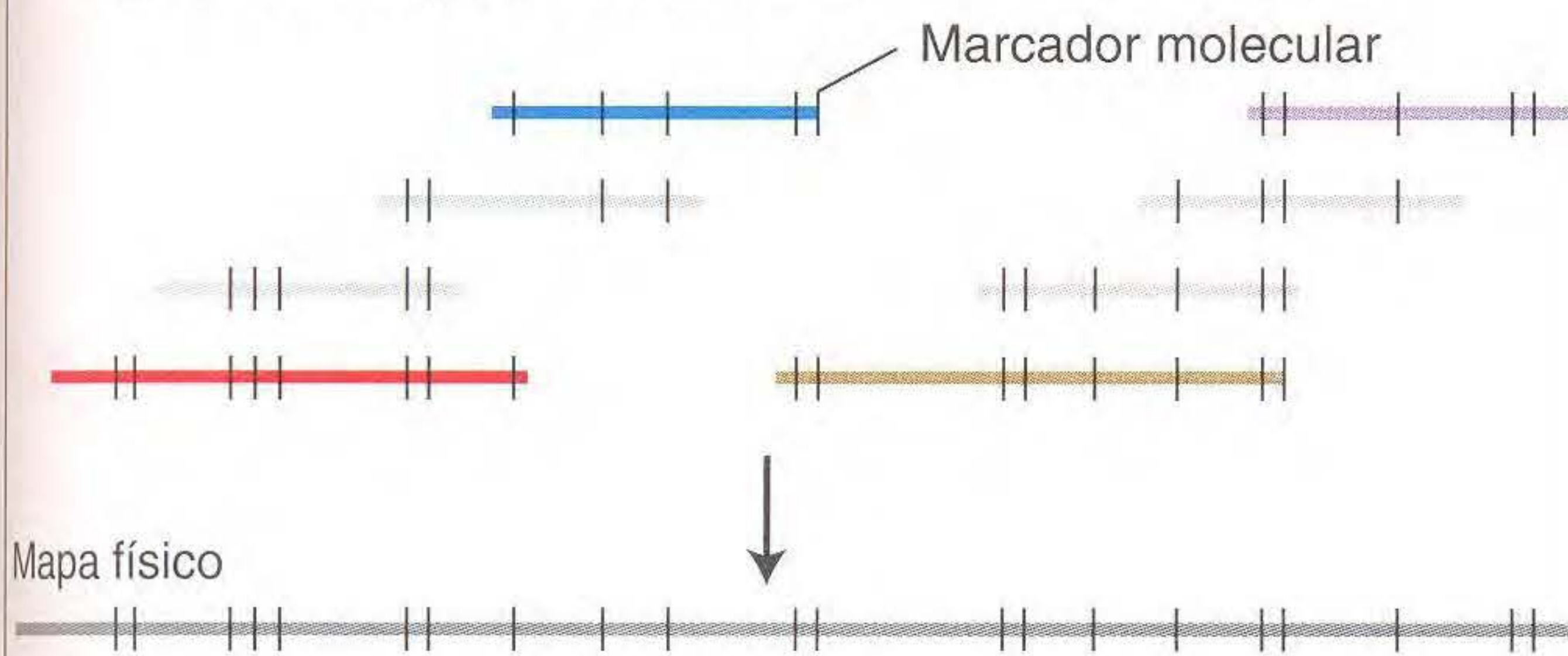
Após ter sido obtido o mapa físico, a etapa seguinte é escolher, de todos os clones usados para construir o mapa, um conjunto de clones minimamente superpostos que, juntos, cobrem todo o genoma (Figura 13.8). Esses clones são então totalmente seqüenciados tratando cada clone genômico como um projeto de seqüenciamento de minigenoma, no qual múltiplas leituras de seqüenciamento para o clone são colocadas juntas usando a lógica do enfoque shotgun de todo o genoma. Finalmente, as seqüências de clones são reunidas em uma seqüência geral de consenso para o genoma de acordo com a ordem conhecida desses clones no mapa físico. Essa habilidade em se basear no mapa físico para ordenar e orientar as seqüências de clones é uma importante vantagem do enfoque de clone ordenado. Uma segunda vantagem importante é a habilidade de incluir alguns elementos repetitivos. A localização de um elemento repetitivo dentro do clone é inequívoca — uma importante vantagem em criar a seqüência de consenso clone por clone.

Os vetores que podem levar grandes insertos são os mais úteis porque o genoma será quebrado em menos pedaços e existirão menos clones a rastrear. Entretanto, mesmo com o uso de vetores que levam grandes insertos, a criação de um mapa físico é uma tarefa árdua. Até os chamados pequenos genomas contêm grandes quantidades de DNA. Considere, por exemplo, o genoma de 100 Mbp do pequeno nematódeo *Caenorhabditis elegans*. Dois vetores comumente usados para levar clones são os cosmídios, que são híbridos de DNA de fago λ e um DNA de plasmídeo bacteriano em uma forma circular, e os vetores BAC derivados do plasmídeo bacteriano F. Como um inserto de um cosmídio médio tem cerca de 40 kb, pelo menos 2.500 cosmídios seriam necessários para cobrir esse genoma, e muitos mais para se certificar de que

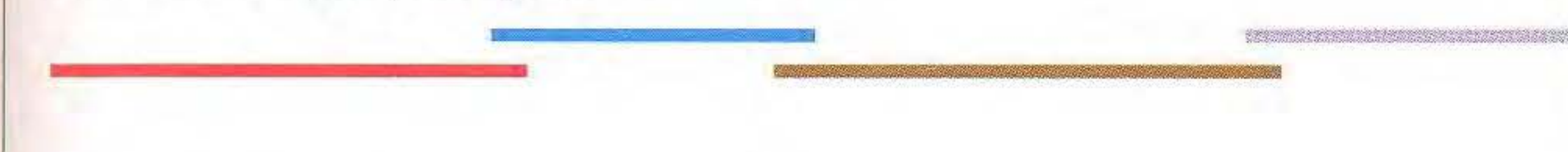
Estratégia para seqüenciamento de clone ordenado

1 Ordenar clones de grandes insertos

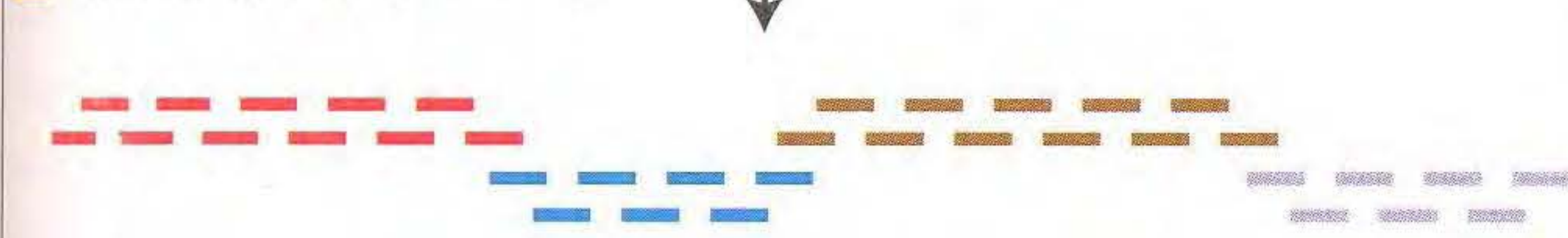
superpondo fingerprints para criar um mapa físico.



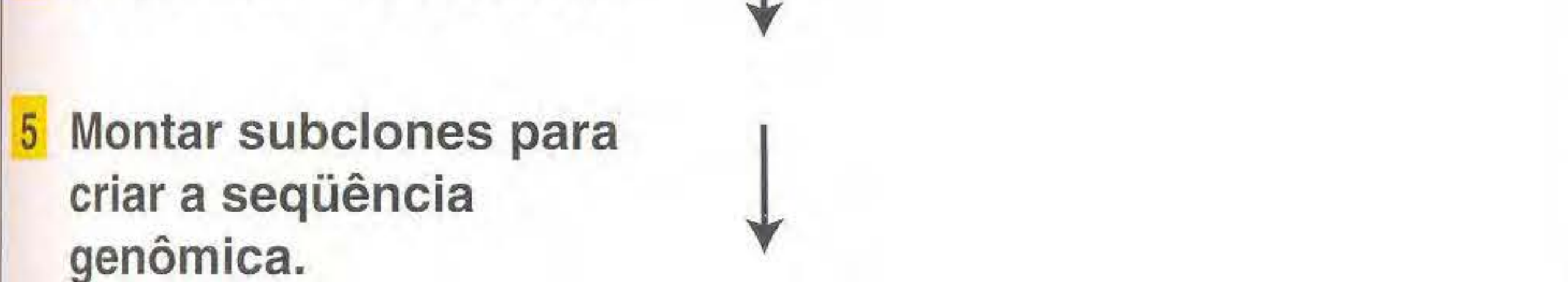
2 Selecionar clones com mínima superposição.



3 Dividir em subclones.



4 Seqüenciar subclones.



5 Montar subclones para criar a seqüência genômica.



FIG. 13.8 O mapeamento físico identifica uma série de clones que se superpõem minimamente. Os clones são divididos em subclones, que são seqüenciados e reordenados.

todos os segmentos do genoma foram representados. Uma biblioteca BAC de *C. elegans* com um sítio médio de inserto de 200 kb simplificaria a tarefa em cinco vezes.

Os dois enfoques básicos para o seqüenciamento do genoma são o seqüenciamento shotgun de todo o genoma e o seqüenciamento de clone ordenado de mapas físicos.

Preenchendo espaços na seqüência

Tanto no shotgun de genoma inteiro quanto no seqüenciamento de clone ordenado, geralmente restam alguns espaços. Espaços ocasionais são encontrados sempre que uma região do genoma por acaso não é encontrada na biblioteca de shotgun — por exemplo, alguns fragmentos de DNA não se replicam bem em vetores de clonagem particulares. Técnicas especiais devem ser usadas para preencher esses espaços nas montagens de seqüências. Se os espaços forem curtos, os fragmentos que faltam podem ser gerados usando as seqüências conhecidas nos finais das montagens como primers para amplificar e analisar a seqüência genômica mediana. Se os espaços forem maiores, podem ser feitas tentativas para clonar essas seqüências em um hospedeiro diferente, tal como leveduras. Se a clonagem em um hospedeiro diferente falha, então os espaços na seqüência podem permanecer.

Se o genoma é seqüenciado pelos padrões “rascunho” ou “terminado” é um julgamento de custo-benefício. É relativamente fácil criar um rascunho, mas muito difícil criar uma seqüência terminada pelos métodos atuais.

► 13.3 Bioinformática: Significado da Seqüência Genômica

A seqüência genômica é um código altamente criptografado contendo a informação para construir e manter um organismo funcional. O estudo do conteúdo de informação dos genomas é chamado de **Bioinformática**. Estamos longe de sermos capazes de ler essa informação do começo ao fim do modo como poderíamos ler um livro. Muito embora saibamos que trincas codificam os aminoácidos nos segmentos codificantes de proteínas, grande parte da informação contida em um genoma não é decifrável por mera inspeção.

A natureza do conteúdo de informação do DNA

O DNA contém informação, mas de que modo é codificada? Convencionalmente, a informação é tida como a soma de todos os produtos gênicos, proteínas e RNA. Entretanto, o conteúdo de informações do genoma é mais complexo que isso. O genoma também contém sítios de ligação para proteínas diferentes e RNA. Muitas proteínas ligam-se a sítios situados no próprio DNA, enquanto outras proteínas e RNA ligam-se a sítios situados no mRNA (Figura 13.9). A seqüência e posições relativas dos sítios permitem que os genes sejam transcritos, recompostos e traduzidos apropriadamente, na época apropriada e no tecido apropriado. Por exemplo, os sítios de ligação de proteínas regulatórias determinam quando, onde e em que nível um gene deve ser expresso. Ao nível do RNA nos eucariontes, os locais dos sítios de ligação para os RNA e proteínas dos spliceossomos determinarão os sítios de corte em 5' e 3' onde os íntrons são removidos. Independentemente de onde um sítio de ligação de fato funcione tal como no DNA ou RNA, o sítio deve ser codificado no DNA. A informação no genoma pode ser tida como a soma de todas as seqüências que codificam proteínas e RNA, mais os sítios de ligação que controlam a época e o local de suas ações. O objetivo principal após a montagem de rascunho do genoma e o objetivo contínuo de como o rascunho é melhorado é a identificação de todos os elementos funcionais do genoma. Esse processo é chamado de **anotação**.

Deduzindo os genes codificantes de proteína da seqüência genômica

Como as proteínas presentes em uma célula determinam amplamente sua morfologia e propriedades fisiológicas, uma das primeiras ordens na análise do genoma e anotação é tentar determinar um inventário de todos os polipeptídeos codificados pelo genoma de um organismo. Esse inventário é chamado de **proteoma** do organismo. Ele pode ser considerado como uma “lista parcial” para a célula. Para determinar a lista de polipeptídeos, a seqüência de cada mRNA

O conteúdo de informação do genoma inclui sítios de ligação

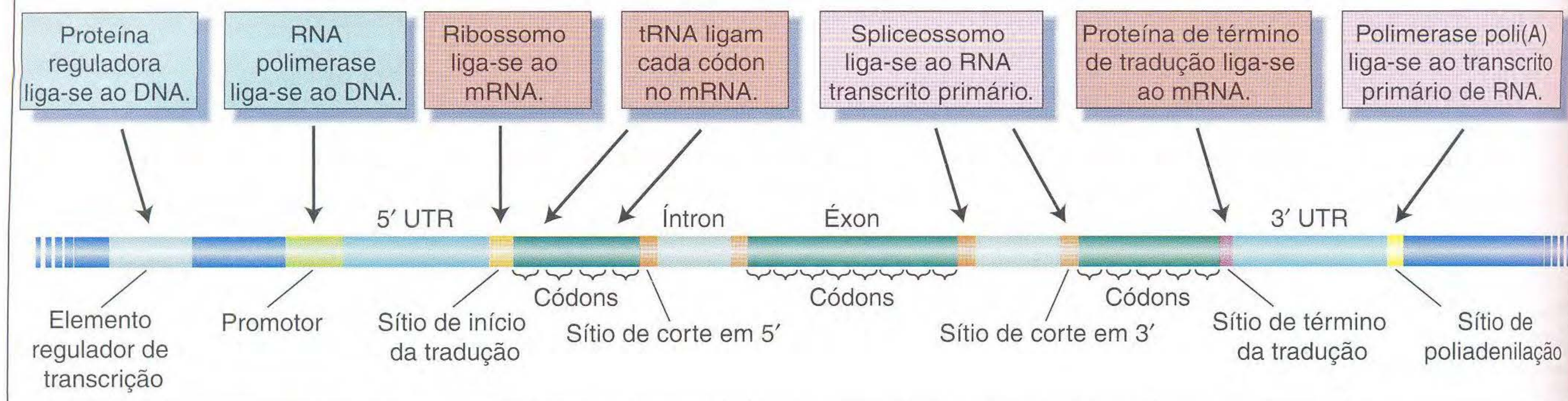


FIG. 13.9 Um gene dentro do DNA pode ser visto como uma série de sítios de ligação para proteínas e RNA.

codificada pelo genoma deve ser deduzida. Devido à remoção dos íntrons, essa tarefa é particularmente desafiadora em eucariontes multicelulares, nos quais os íntrons são a norma. Em humanos, por exemplo, um gene médio tem cerca de 10 éxons. Além disso, muitos genes codificam éxons alternativos; isto é, alguns éxons são incluídos em algumas versões de um mRNA processado, mas não são incluídos em outras (veja o Capítulo 8). Os mRNA processados alternativamente podem codificar polipeptídeos tendo muitas, mas não todas, as seqüências de aminoácidos em comum. Muito embora tenhamos muitos exemplos de genes completamente seqüenciados e mRNA, ainda não podemos identificar os sítios de corte 5' e 3' apenas pela seqüência de DNA com um alto grau de precisão. Portanto, não podemos estar certos de quais seqüências são íntrons. As previsões de éxons alternativamente usados são ainda mais propensas a erros. Por muitos motivos, deduzir todas as partes de polipeptídeos em eucariontes superiores é um grande problema. Alguns enfoques são os seguintes.

Deteção de ORF O principal enfoque para produzir uma lista de polipeptídeos é o uso da análise computacional da seqüência do genoma para prever um mRNA e seqüências de polipeptídeos, uma parte importante da Bioinformática. O enfoque básico é procurar seqüências que tenham as características de genes. Essas seqüências teriam tamanhos de genes e seriam compostas de códons com sentido após possíveis íntrons terem sido removidos. As seqüências apropriadas das pontas 5' e 3' estariam presentes, tais como códons de início e fim. As seqüências com essas características típicas de genes são chamadas de **matrizes de leitura abertas (ORF)**. Para encontrar candidatas à ORF, o computador faz uma varredura da seqüência do DNA em ambos os filamentos em cada matriz de leitura. Como existem três possíveis matrizes de leitura em cada filamento, existem seis matrizes de leitura possíveis no total.

Evidência direta de seqüências de cDNA Outro meio de identificar ORF e éxons é pela análise da expressão de mRNA. Essa análise é feita criando-se bibliotecas de moléculas de DNA que são complementares a seqüências de mRNA, chamadas cDNA. As seqüências complementares de DNA são

extremamente valiosas de dois modos. Primeiro, elas são a evidência direta de que determinado segmento do genoma é expresso e pode, portanto, codificar um gene. Segundo, como cDNA é complementar ao mRNA final, os íntrons dos transcritos primários foram removidos, o que facilita muito a identificação dos éxons e íntrons de um gene (Figura 13.10). O alinhamento dos cDNA com sua seqüência genômica correspondente claramente delineia os éxons, e, assim, os íntrons são revelados à medida que as regiões caem entre os éxons. No cDNA, a ORF deve ser contínua do início do códon até o códon de fim. Assim, as seqüências de cDNA são muito ajudadas na identificação da leitura da matriz correta, incluindo os códons de início e fim. Evidências de todo o cDNA são obtidas como a prova padrão principal de que foi identificada a seqüência de uma unidade de transcrição, incluindo seus éxons e sua localização no genoma.

Além do tamanho total das seqüências de cDNA, existem grandes conjuntos de dados de cDNA para os quais apenas as pontas 5' ou 3' ou ambas foram seqüenciadas. Essas seqüências curtas de cDNA são chamadas de **seqüências expressas marcadas (EST)**. Essas seqüências podem ser alinhadas com o DNA genômico e, assim, usadas para determinar as pontas 5' e 3' dos transcritos, isto é, para determinar os limites do transcrito como mostrado na Figura 13.10.

Previsões dos sítios de ligação Como já foi discutido, um gene consiste em um segmento de DNA que codifica um transcrito, bem como em sinais reguladores que determinam quando, onde e como a grande parte desse transcrito é feita. Por sua vez, esse transcrito tem os sinais necessários para determinar sua recomposição em mRNA e a tradução desse mRNA em um polipeptídeo (Figura 13.11). Existem hoje programas de computador estatísticos para “encontrar genes” que procuram as seqüências previstas dos vários sítios de ligação usados como promotores, para a transcrição dos sítios de início, os sítios de corte de 3' e 5', e para os códons de início de tradução dentro do DNA genômico. Essas previsões são baseadas em motivos de consenso para tais seqüências conhecidas, mas elas não são perfeitas.

Uso de polipeptídeo e similaridade de DNA Como os organismos têm ancestrais comuns, eles também têm muitos genes com seqüências similares em comum. Assim, um gene provavelmente

cDNA e EST revelam éxons ou finais de genes nas pesquisas genômicas

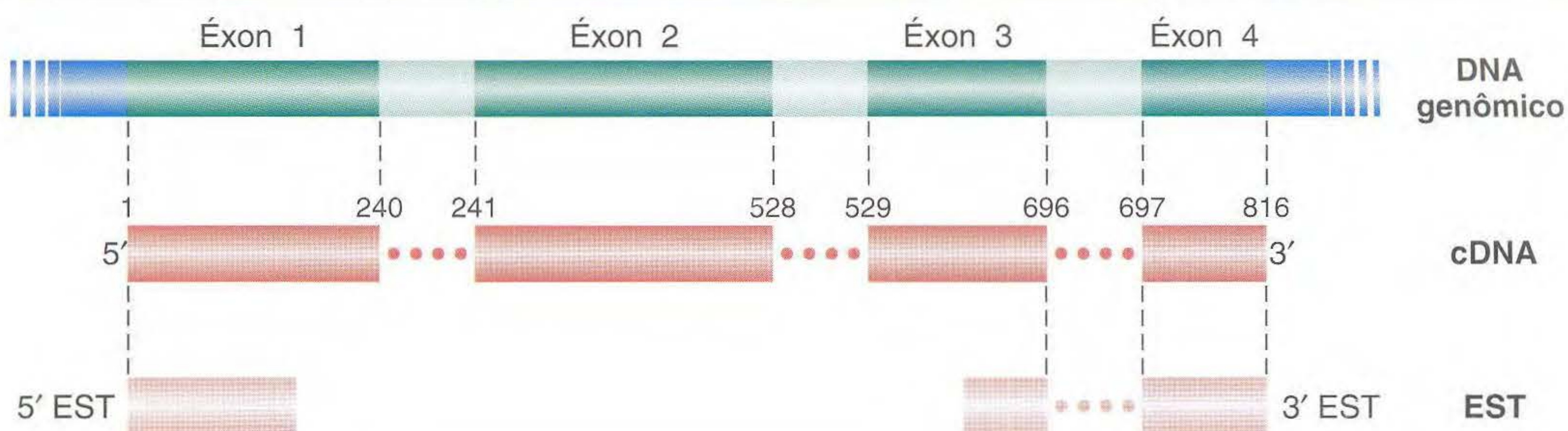


FIG. 13.10 Alinhamento de DNA (cDNA) complementares totalmente seqüenciados e seqüências marcadas expressas (EST) com DNA genômico. As linhas tracejadas indicam regiões de alinhamento; para o cDNA, essas regiões são os éxons do gene. Os pontos entre os segmentos de cDNA ou EST indicam regiões do DNA genômico que não se alinham com cDNA ou seqüências EST; essas regiões são locais de íntrons. Os números acima da linha de cDNA indicam as coordenadas de bases da seqüência do cDNA, onde a base 1 está mais para 5' e a base 816 está mais para 3' do cDNA. Para as EST, apenas uma curta leitura de seqüência é obtida de cada ponta (5' e 3') do cDNA correspondente. Essas leituras de seqüências estabelecem os limites da unidade de transcrição, mas não são informativas da estrutura interna do transcrito, a menos que seqüências EST cruzem um íntron (como é verdade para 3' EST mostrada aqui).

Caçadas genômicas para vários sítios de ligação

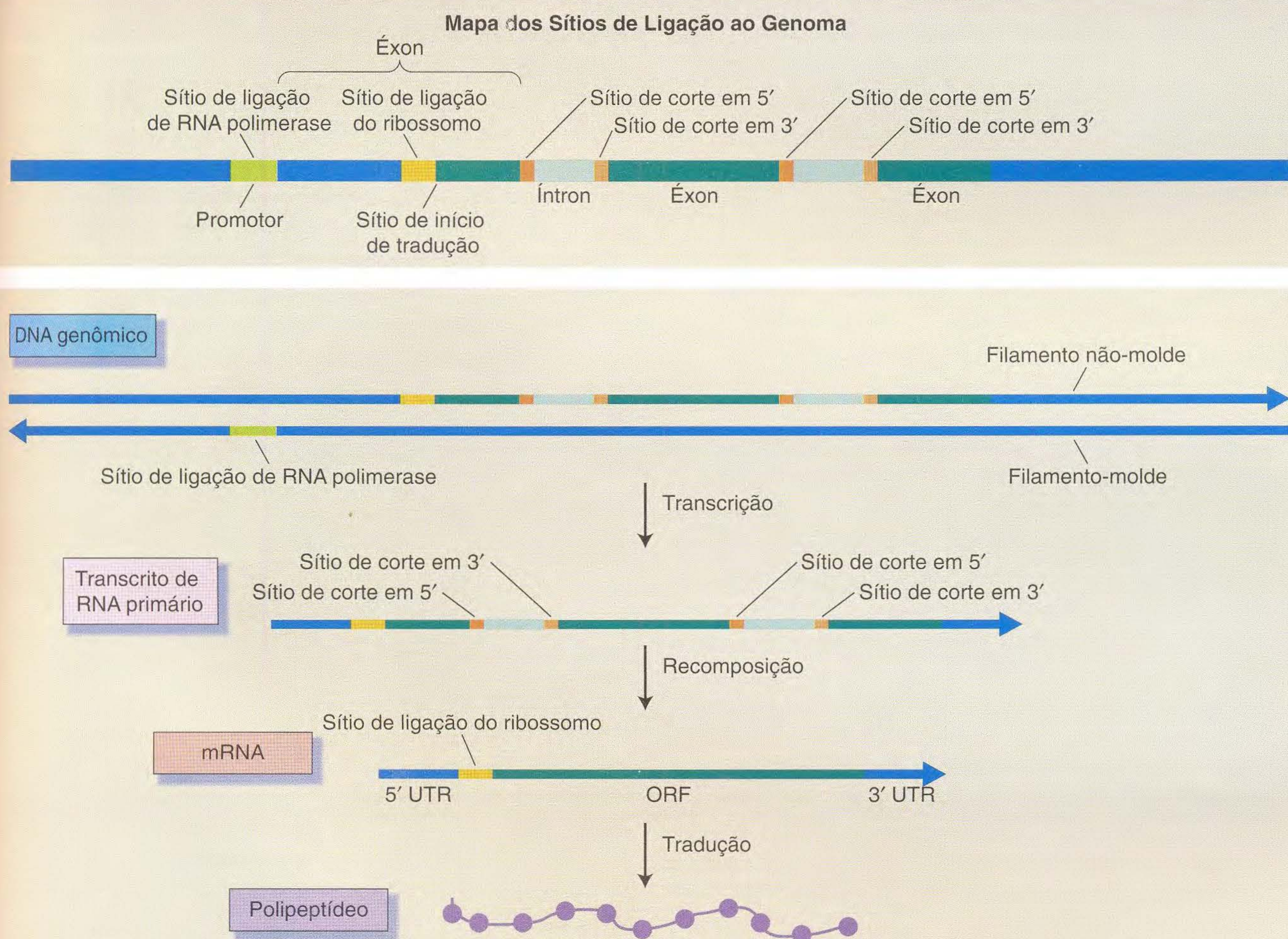


FIG. 13.11 Transferência de informação eucariótica do gene para a cadeia polipeptídica. Note que os "sítios de ligação" de DNA e RNA que estão ligados a complexos proteicos iniciam os eventos de transcrição, recomposição e tradução.

terá correlatos entre os genes isolados e seqüenciados em outros organismos, especialmente nos muito correlatos. Os genes candidatos previstos pelas técnicas precedentes muitas vezes podem ser verificados por comparação com todas as outras seqüências de genes encontradas. Uma seqüência candidata é submetida como uma “seqüência indagada” aos bancos de dados públicos contendo um registro de todas as seqüências de genes conhecidas. Esse procedimento é chamado de pesquisa BLAST (de Basic Local Alignment Search Tool). A seqüência pode ser submetida como uma seqüência de nucleotídeos (pesquisa BLASTn) ou como uma seqüência traduzida de aminoácidos (BLASTp). O computador pesquisa o banco de dados e retorna uma lista parcial ou total de “alvos”, começando com os que mais correspondem. Se a seqüência candidata assemelha-se muito com um gene previamente identificado de outro organismo, então essa semelhança fornece uma forte indicação de que o gene candidato é de fato um gene. Equivalências menos próximas também são úteis. Por exemplo, uma identidade de aminoácidos apenas de 35%, mas em posições idênticas, é um forte indicador de que duas proteínas têm uma estrutura tridimensional comum.

As pesquisas BLAST são usadas de muitos outros modos, mas a meta é sempre descobrir mais sobre alguma seqüência identificada de interesse.

Previsões baseadas em tendenciosidades de códon Lembre-se, do Capítulo 9, que o código triplo para aminoácidos é redundante, isto é, a maioria dos aminoácidos é codificada por dois ou mais códons (veja Figura 9.6). Os múltiplos códons para um único aminoácido são chamados de *códons sinônimos*. Em determinada espécie, nem todos os códons sinônimos para um aminoácido são usados com igual freqüência. Alguns códons estão presentes muito mais freqüentemente nos mRNA (e, assim, no DNA que os codifica). Por exemplo, em *D. melanogaster*, dos dois códons para cisteína, UGC é usado em 73% do tempo, enquanto UGU é usado em 27%. Esse uso é um diagnóstico de *Drosophila* porque, em outros organismos, esse padrão de “tendenciosidade de códons” é bem diferente. As tendenciosidades de códon são consideradas devidas à abundância relativa dos tRNA complementares a esses vários códons em determinada espécie. Se o uso do códon de uma ORF prevista corresponde ao padrão conhecido de uso da espécie, então essa correspondência é uma evidência de apoio de que a ORF proposta é genuína.

Juntamente Um resumo de como fontes diferentes de informação são combinadas para criar o melhor mRNA possível e as previsões gênicas é mostrado na Figura 13.12. Esses

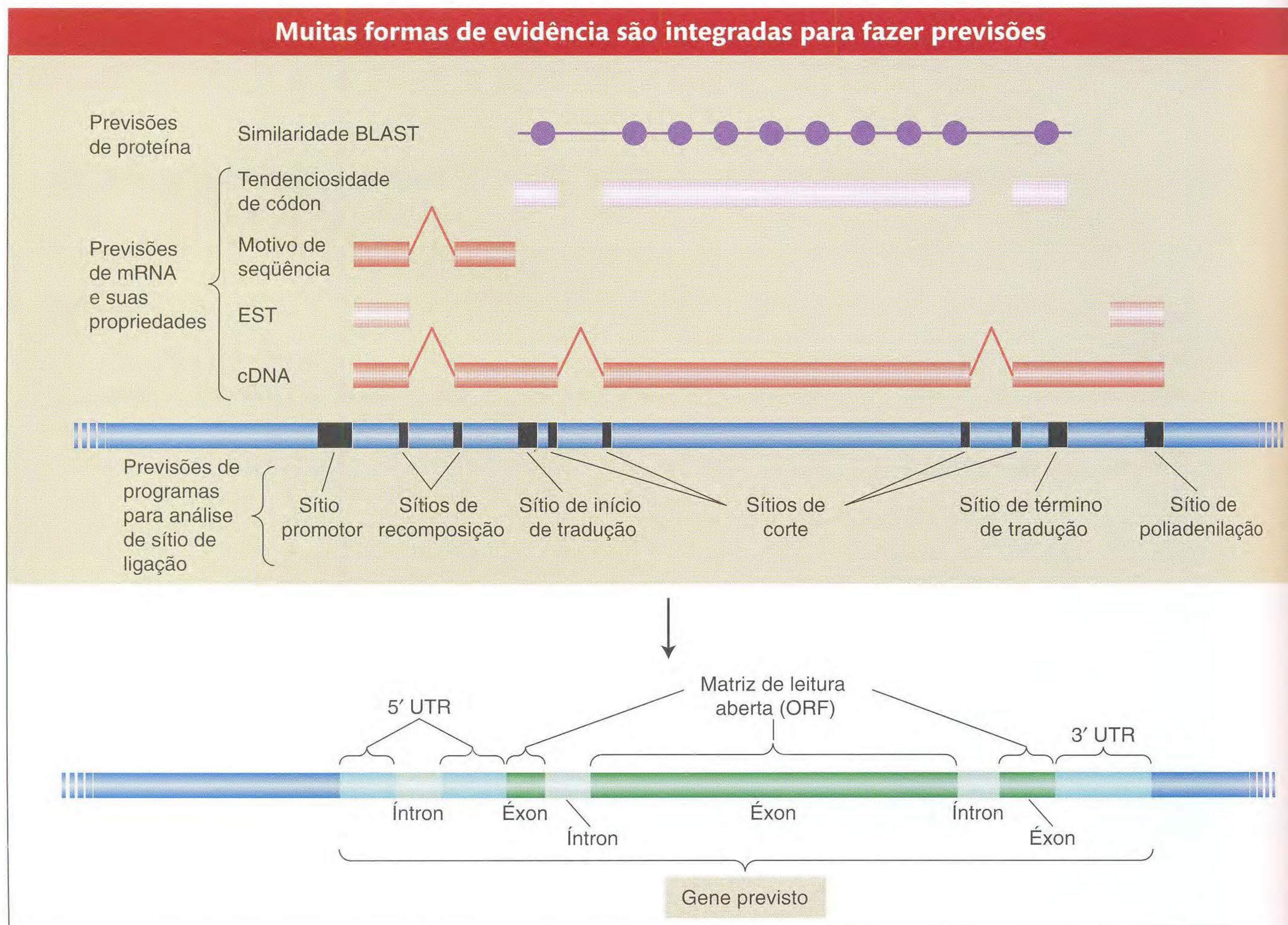


FIG. 13.12 As diferentes formas de evidenciar produtos gênicos – cDNA, EST, eventos similares de BLAST, tendenciosidade de códon e motivos – são integradas para fazer previsões gênicas. Onde várias classes de evidências são encontradas associadas a determinada seqüência genômica do DNA, há uma confiança maior na probabilidade de que a previsão de um gene seja precisa.

tipos diferentes de evidências são complementares e podem fazer uma validade cruzada uma da outra. Por exemplo, a estrutura de um gene pode ser deduzida da evidência da similaridade de proteína dentro de uma região do DNA genômico ligado pelas EST 5' e 3'. São possíveis previsões úteis mesmo sem uma sequência de cDNA ou evidência de similaridades de proteína. Um programa de previsão de sítio de ligação pode propor uma ORF hipotética, e a tendenciosidade apropriada do códon seria a evidência de apoio.

Previsões do mRNA e estrutura de polipeptídeos da sequência de DNA genômico dependem da integração da informação da sequência de cDNA, previsões de ligação, similaridades de polipeptídeos e tendenciosidade de códons.

Consideremos alguns dos *insights* de nossa primeira visão das estruturas genômicas gerais e listas de partes globais de algumas espécies cujos genomas foram seqüenciados. Começaremos com nós mesmos. O que podemos aprender observando o genoma humano em si? Então, veremos o que podemos aprender comparando nosso genoma com outros.

▶ 13.4 A Estrutura do Genoma Humano

Ao descrevermos a estrutura geral do genoma humano, devemos primeiro examinar sua estrutura repetida. Uma fração considerável do genoma humano, cerca de 45%, é repetitiva. Grande parte desse DNA repetitivo é composto de cópias de elementos de transposição. De fato, mesmo na cópia única do DNA restante, uma fração tem seqüências sugerindo que podem descender de antigos elementos de transposição que são imóveis e acumularam mutações aleatórias, fazendo com que divirjam em seqüência dos elementos de transposição ancestrais. Assim, grande parte do genoma humano parece ser composto de “caronas” genéticas.

Apenas uma pequena parte do genoma humano codifica polipeptídeos. Isso é, um pouco menos de 3% codifica éxons *dos mRNA*. Os éxons são tipicamente pequenos (cerca de 150 bases), enquanto os íntrons são grandes, muitos com mais 1.000 bases e alguns estendendo-se a mais 100.000 bases. Os transcritos são compostos de uma média de 10 éxons, embora muitos tenham substancialmente mais. Finalmente, os íntrons podem ser recompostos do mesmo gene em locais que variam. Essa variação no local de sítios de recomposição adiciona considerável diversidade ao mRNA e seqüências polipeptídicas. Com base nos dados atuais de cDNA e EST, 60% dos genes humanos codificantes de proteínas provavelmente têm duas ou mais variantes de corte. Em média, existem cerca de três variantes de corte por gene. Assim, o número de proteínas distintas codificadas pelo genoma humano é cerca de três vezes maior do que o número de genes reconhecidos.

O número de genes no genoma humano não tem sido fácil de quantificar. No rascunho inicial do genoma humano, havia uma estimativa de 30.000 a 40.000 genes codificantes de

proteínas. Entretanto, a arquitetura complexa desses genes e do genoma pode dificultar a anotação. Algumas seqüências registradas como genes na verdade podem ser éxons de genes maiores. Além disso, existem mais de 19.000 **pseudogenes**; estes são ORF ou ORF parciais que podem, a princípio, parecer genes, mas são não-funcionais ou inativos devido ao modo de sua origem ou a mutações. Os chamados **pseudogenes processados** são seqüências de DNA que sofreram transcrição reversa de RNA e foram inseridos aleatoriamente no genoma. Cerca de 90% dos pseudogenes humanos parecem ser desse tipo. Aproximadamente 900 pseudogenes parecem ser genes convencionais que adquiriram uma ou mais mutações perturbadoras de ORF no curso da evolução. Quando os desafios nas anotações foram superados, o número estimado de genes no genoma humano caiu drasticamente. O número atual, de junho de 2006, é de cerca de 22.000 genes, e o número final pode ser próximo de 19.000 genes funcionais.

Um modo pelo qual a anotação do genoma humano progrediu é pelas seqüências terminais de cada cromossomo, um a um. Essas seqüências então se tornaram o campo de busca na caçada por genes candidatos. Um exemplo das previsões dos genes para um cromossomo a partir do genoma humano é mostrado na Figura 13.13. Tais previsões são continuamente revistas à medida que novos dados e novos programas de computação tornam-se disponíveis. O estado atual de previsões pode ser visto em muitos *web sites*, mais notadamente nos bancos de dados públicos de DNA nos EUA e Europa (veja o Apêndice B). Essas previsões são as melhores inferências atuais dos genes codificantes de proteínas presentes nas espécies seqüenciadas e, como tais, são progressos em andamento.

As proteínas podem ser agrupadas em famílias de proteínas correlatas, similares em estrutura e função, com base na similaridade na seqüência de aminoácidos. Para determinada família de proteínas que é conhecida em muitos organismos, o número de proteínas na família é geralmente maior em humanos do que em não-vertebrados, cujos genomas foram seqüenciados. As proteínas são compostas de domínios modulares que são misturados e pareados para desempenhar papéis diferentes. Muitos domínios estão associados a funções biológicas específicas. O número de domínios modulares por proteína também parece ser maior em humanos do que em organismos invertebrados.

À medida que surgem informações mais refinadas sobre o genoma humano, podem ser conhecidas características adicionais. Um exemplo recente é a seqüência de mapa terminada de um dos mais estudados cromossomos humanos, o cromossomo 7. Inicialmente, esse cromossomo foi intensamente estudado porque contém o gene que, quando mutado, causa fibrose cística. A localização do gene de fibrose cística foi identificada nos primeiros dias do Projeto Genoma Humano pela superposição do mapa de ligação ao físico e mapas de seqüenciamento, como descrito no Capítulo 4. Os grupos continuaram a estudar esse cromossomo em detalhes. Cerca de 1.700 genes são conhecidos ou previstos residindo no cromossomo 7. Cerca de 800 clones de mapas físicos foram situados no cromossomo humano 7.

A sequência de mapa do cromossomo humano 20

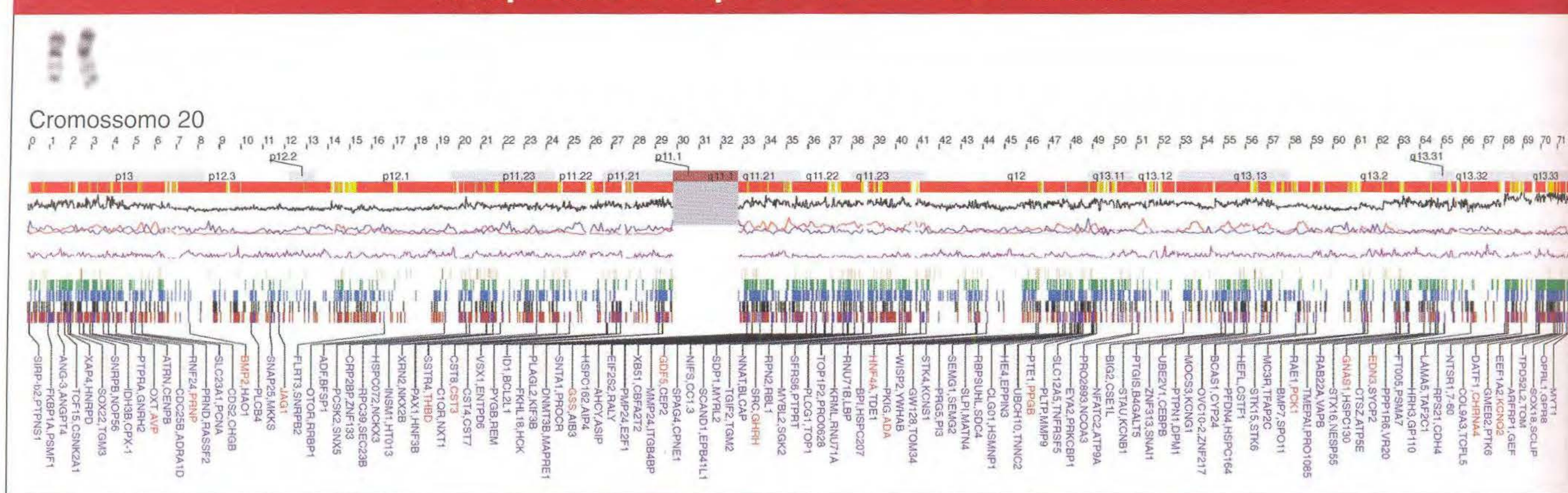


FIG. 13.13 Vários genes foram identificados no cromossomo humano 20. As coordenadas do mapa recombinacional e citogenético são mostradas nas linhas de cima da figura. Vários gráficos mostrando a densidade gênica e as propriedades diferentes do DNA são mostrados nas seções do meio. Os identificadores dos genes previstos são mostrados na parte inferior do painel. [Cortesia de Jim Kent, Ewan Birney, Darryl Leja, and Francis Collins. Adaptado do International Human Genome Sequencing Consortium, "Initial Sequencing and Analysis of the Human Genome," *Nature* 409, 2001, 860-921.]

Um uso dos clones de mapas físicos é no mapeamento de rearranjos de pontos de quebra associados a doença humana. Os rearranjos cromossômicos são uma classe de mutações que resultam da quebra de um cromossomo em um locus, o ponto de quebra do rearranjo e sua reunião com outro sítio cortado similar no mesmo cromossomo ou em um diferente. Essas quebras causam mutações quando um gene reside no ponto de quebra. Com o uso de clones físicos, cerca de 1.600 pontos de quebra de rearranjos associados a doenças humanas foram mapeados no cromossomo 7, criando um mapa citogenético de alta densidade (Figura 13.14). Desses pontos de quebra, 440 foram seqüenciados, permitindo a associação de fenótipos mutantes e genes encontrados nas seqüências de DNA.

► 13.5 Genômica Comparativa

Um dos meios mais poderosos para avançar a análise do nosso genoma e de outros é a comparação da estrutura do genoma e a seqüência entre espécies correlatas. Como a seleção natural em geral rejeita mutações que diminuam a adaptabilidade, os genes e outras seqüências funcionais de DNA são conservados durante longos períodos da evolução. Uma via de seqüência de DNA que é comum em espécies divergentes provavelmente desempenha uma função necessária, e tais vias comuns podem ser usadas para guiar os estudos destinados a descobrir essas funções. Além disso, os genes identificados em uma espécie-modelo provavelmente são identificáveis, com base na seqüência e localização do genoma, em espécies correlatas.

Além da identificação das regiões conservadas, a **genômica comparativa** tem o potencial de revelar como as espécies divergem. As espécies evoluem e as características mudam por alterações na seqüência de DNA. As comparações entre espécies de genomas podem revelar eventos únicos a determinadas linhagens que podem contribuir para diferenças na fisiologia, comportamento ou anatomia. Aqui veremos

alguns exemplos de como a genômica comparativa revela o que é similar e diferente entre as espécies.

Camundongos e humanos

A seqüência do genoma de camundongos tem sido particularmente informativa para compreender o genoma humano, devido ao papel do camundongo há muito tempo como um modelo de espécie genética, ao vasto conhecimento de sua genética clássica e à relação evolutiva do camundongo com os humanos. As linhagens humanas e de camundongos divergiram há aproximadamente 75 milhões de anos, o que é tempo suficiente para as mutações fazerem seus genomas diferirem, em média, em cerca de um a cada dois nucleotídeos. Assim, as seqüências comuns aos genomas de camundongos e humanos provavelmente indicam funções comuns.

A primeira etapa em comparar genomas é a identificação dos genes mais proximamente relacionados, chamados de **homólogos**. É importante distinguir aqui duas classes de genes homólogos. Alguns homólogos estão no mesmo locus genético herdado de um ancestral comum; esses genes são chamados de **ortólogos**. Entretanto, muitos genes pertencem a famílias que aumentaram (e diminuíram) em número no curso da evolução. Os genes que são relacionados por eventos de duplicação gênica em um genoma são chamados de **parálogos**. Nas comparações genômicas, nem sempre é possível identificar a correlação entre parálogos em famílias de genes.

Os homólogos são identificados porque têm seqüências de DNA similares. A análise do genoma de camundongo indica que o número de genes codificantes de proteínas que ele contém é similar ao do genoma humano. Uma maior inspeção dos genes de camundongo revela que 99% de todos os genes de camundongo têm algum homólogo no genoma humano, e que pelo menos 99% de todos os genes humanos têm algum homólogo no genoma de camundongo. Assim, os tipos de proteínas codificadas em cada genoma são os mesmos. Além

Mapa citogenético do cromossomo humano 7

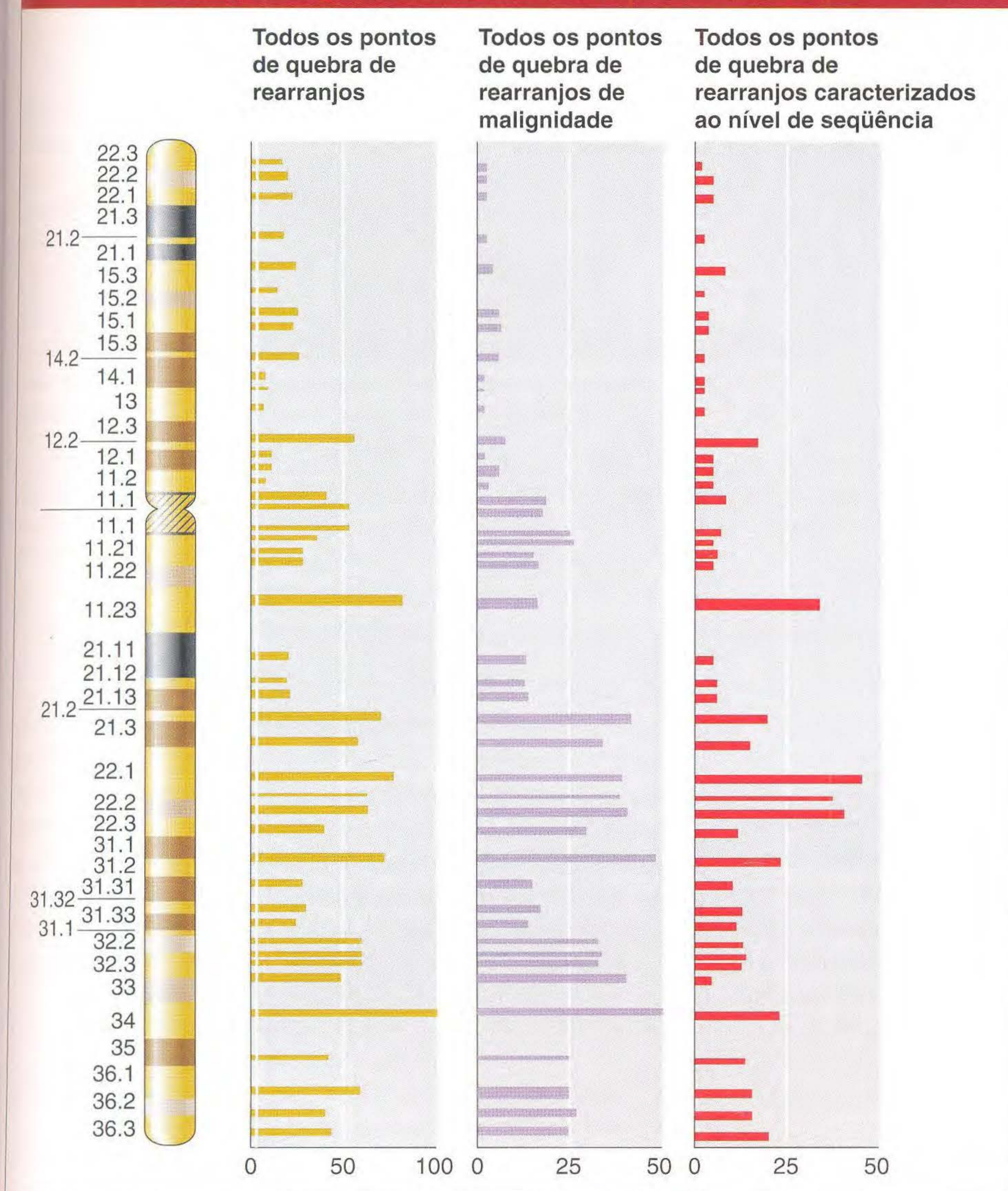


FIG. 13.14 Pontos de quebra de rearranjos de pacientes com distúrbios genéticos foram mapeados no cromossomo 7, criando um mapa citogenético. [De W. S. Scherer et al., "Human Chromosome 7: DNA Sequence and Biology," *Science* 300, 2003, 769 and 771, Figs. 2 and 5.]

disso, cerca de 80% de todos os genes de camundongos e humanos são claramente ortólogos identificáveis.

As similaridades entre os genomas vão bem além do inventário de genes codificantes de proteínas para a organização genômica geral. Mais de 90% dos genomas humano e de camundongos podem ser repartidos em regiões correspondentes de **sintenia** conservada, onde a ordem dos genes dentro de blocos de tamanhos variados é a mesma que a ordem nos ancestrais comuns mais recentes das duas espécies. Essa sintenia é muito útil na correlação dos mapas dos dois genomas. Por exemplo, o cromossomo humano 17 é ortólogo a um único cromossomo de camundongo (cromossomo 11). Embora tenham ocorrido amplos rearranjos intracromossômicos no cromossomo humano, existem 23 segmentos de sequências colineares com mais 100 kb de tamanho (Figura 13.15).

Existem algumas diferenças detectáveis entre os inventários de genes de camundongos e genes humanos. O camundongo possui mais cópias de genes que tomam parte na imunidade, olfação e reprodução. O número ampliado de

tais genes pode sugerir que esses sistemas fisiológicos evoluíram na linhagem dos roedores. Isso certamente faz sentido com relação aos estilos de vida das espécies. Os camundongos dependem muito de seu sentido de olfato, e encontram constelações muito diferentes de patógenos do que os encontrados pelos humanos (ou nossos ancestrais macacos). Além disso, nossas fisiologias reprodutivas são amplamente diferentes. Ademais, essas diferenças em conteúdo gênico são relativamente modestas à luz das grandes diferenças em anatomia e comportamento. A similaridade geral nos genomas de camundongos e humanos corresponde ao quadro do exame das ferramentas genéticas que controlam o desenvolvimento em espécies diferentes (veja o Capítulo 12) — as grandes diferenças podem evoluir de genomas contendo conjuntos similares de genes.

Os genomas de mamíferos, inclusive o dos humanos, contêm conjuntos similares de genes.

Os genomas humano e de camundongo têm grandes blocos sintênicos de genes em comum

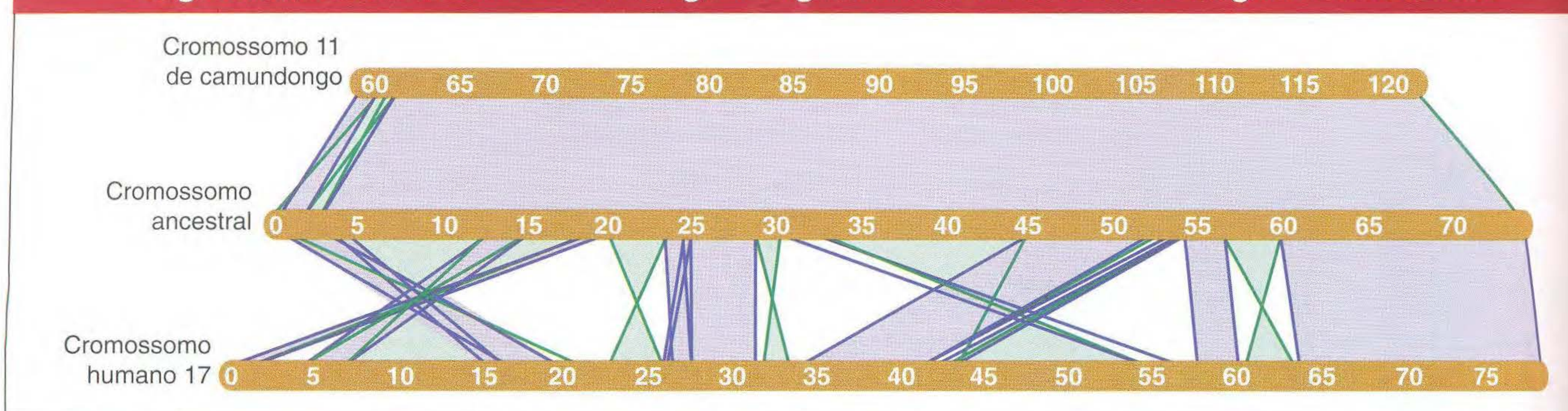


FIG. 13.15 Sintenia entre o cromossomo humano 17 e o cromossomo 11 de camundongo. São mostrados grandes blocos sintênicos conservados de 100 kb ou maiores no cromossomo humano 17, cromossomo 11 de camundongo e o cromossomo deduzido de seu último ancestral comum (reconstruído por análise de outros genomas de mamíferos). Blocos diretos de sintenia são mostrados em azul-claro; blocos invertidos são mostrados em verde. Os tamanhos dos cromossomos são indicados em megabases (Mb). [De M.C. Zody et al., "DNA Sequence of Human Chromosome 17 and Analysis of Rearrangement in the Human Lineage," *Nature* 440, 2006, 1045-1049, Fig. 2.]

Esse mesmo tema é ilustrado comparando nosso genoma com o de nosso parente mais próximo, o chimpanzé.

Genômica comparativa de chimpanzés e humanos

Os chimpanzés e humanos tiveram um ancestral comum há cerca de 6 milhões de anos. Desde essa época, as diferenças genéticas se acumularam por mutações que ocorreram em cada linhagem. O sequenciamento genômico revelou que existem cerca de 35 milhões de diferenças de um só nucleotídeo entre chimpanzés e humanos, correspondendo a cerca de 1,06% de grau de divergência. Além disso, cerca de 5 milhões de inserções ou deleções, desde um único nucleotídeo até mais de 15 kb, contribuem com um total de cerca de 90 Mb de sequência de DNA divergente (cerca de 3% do genoma geral). A maioria dessas inserções ou deleções ficam fora das regiões codificantes.

No geral, as proteínas codificadas pelos genomas humano e de chimpanzé são extremamente similares. Vinte e cinco por cento de todas as proteínas ortólogas são idênticas em sequência. A maioria das proteínas diferem apenas por cerca de duas substituições de aminoácidos. Existem algumas diferenças detectáveis entre os chimpanzés e os humanos nos conjuntos de genes funcionais. Cerca de 80 dos genes que eram funcionais em seus ancestrais comuns não funcionam mais nos humanos, devido à sua deleção ou ao acúmulo de mutações. Algumas dessas mudanças podem contribuir para diferenças em fisiologia.

Além das mudanças em determinados genes, as duplicações de segmentos cromossômicos em uma única linhagem contribuíram para a divergência genômica. Mais de 170 genes no genoma humano e mais de 90 genes no genoma de chimpanzé estão presentes em grandes segmentos duplicados. Essas duplicações são responsáveis por uma maior quantidade da divergência genômica total do que todas as mutações de nucleotídeos únicos combinadas. Entretanto, ainda não está claro se elas contribuem para importantes diferenças.

◆ O QUE OS GENETICISTAS ESTÃO FAZENDO HOJE

Elementos conservados e ultraconservados não-codificantes

A discussão até agora enfocou unicamente as regiões codificantes de proteínas do genoma. Essa ênfase é devida mais à facilidade analítica do que à importância biológica. Devido à simplicidade e universalidade do código genético, a detecção de ORF e éxons é bem mais fácil do que a detecção das sequências funcionais não-codificantes. Como já foi dito, apenas 3% do genoma humano codificam éxons de mRNA, e menos de metade dessas sequências éxon, cerca de 1 a 2% do total do DNA do genoma, codificam sequências de proteínas. Assim, mais de 98% de nosso genoma não codificam proteínas. Como identificamos outras partes funcionais do genoma?

Além das regiões promotoras de genes, que contêm motivos de sequências típicas (veja o Capítulo 11), é difícil atribuir uma função a tantas sequências não-codificantes. Entretanto, um modo para situar elementos não-codificantes potencialmente funcionais é procurar sequências conservadas, que não mudaram muito em milhões de anos de evolução. As comparações de genomas de camundongos e humanos revela que cerca de 5% de todas as sequências são conservadas, com cerca de um terço dessa quantidade consistindo em sequências codificantes de proteínas, e os dois terços restantes consistindo em sequências que não codificam proteínas. Assim, a proporção do genoma que controla como nossos genes são regulados pode ser maior do que a que codifica proteínas.

A identificação dos elementos não-codificantes funcionais pode ser feita pela genômica comparativa. Por exemplo, podemos procurar sequências altamente conservadas de tamanho modesto entre algumas espécies, ou sequências menos perfeitamente conservadas de maior tamanho entre um número maior de espécies. As comparações dos genomas humano, de rato e de camundongos levaram à identificação dos chamados *elementos ultraconservados*, que são sequências perfeitamente conservadas entre as três espécies. As sequên-

cias desses genomas encontraram mais de 5.000 seqüências de mais de 100 pb e 481 seqüências de mais de 200 pb que são absolutamente conservadas.

Ampliando essa análise para incluir o genoma de cachorro, os pesquisadores encontraram mais de 140.000 elementos altamente conservados com 50 pb ou mais de comprimento fora das seqüências codificantes de proteínas. Embora 50% desses elementos sejam encontrados em regiões pobres em genes, eles são mais ricamente concentrados perto de importantes genes reguladores para o desenvolvimento. A maioria dos elementos altamente conservados não-codificantes pode amplamente tomar parte na regulação da expressão das ferramentas genéticas para o desenvolvimento de mamíferos e outros vertebrados.

Como podemos verificar que tais elementos conservados têm um papel na regulação gênica? Esses elementos podem ser testados do mesmo modo que os elementos reguladores de ação cis transcricionais examinados em capítulos anteriores, com o uso de genes repórter. Um pesquisador coloca regiões reguladoras candidatas adjacentes a um promotor e gene repórter, e introduz o gene repórter em espécies hospedeiras. Um desses exemplos é mostrado na Figura 13.16. Um elemento que é altamente conservado entre mamíferos, galinhas e espécies de sapo fica a 488 kb da ponta 3' do gene humano *ISL1*, que codifica uma proteína necessária para a diferenciação do neurônio motor. Esse elemento foi colocado antecedente a um promotor e ao gene repórter de β -galactosidase (*lacZ*), e o conjunto foi injetado nos pronúcleos de ovócitos fertilizados de camundongo. O gene repórter é então visto sendo expresso ao longo da coluna dorsal e na cabeça, como seria esperado para a localização de futuros neurônios motores (veja Figura 13.16). Mais significativamente, o padrão de expressão corresponde a uma parte do padrão de expressão do gene *ISL1* nativo de camundongo (supostamente, outros elementos não-codificantes controlam outras características da expressão de *ISL1*). Muitos milhares de elementos reguladores não-codificantes humanos podem provavelmente ser identificados com base na conservação da seqüência e em sua atividade nas triagens de repórter.

As comparações dos proteomas de camundongo, humanos e chimpanzés, bem como a identificação e análise de seus elementos comuns não-codificantes, destacam a natureza conservativa da evolução do genoma. Entretanto, algumas diferenças marcantes no conteúdo do genoma foram reveladas pela genômica comparativa, com grandes implicações para a medicina humana. Veremos um desses casos em seguida.

Genômica comparativa de *E. coli* não-patogênica e patogênica

A *Escherichia coli* é encontrada em grande número em nossa boca e em nosso trato intestinal, e essa espécie é geralmente um simbionte benigno. Devido a seu papel central nas pesquisas genéticas, foi um dos primeiros genomas bacterianos seqüenciados. No genoma de *E. coli*, há cerca de 4,6 Mb de tamanho e existem 4.405 genes. Entretanto, chamá-lo de “o genoma de *E. coli*” de fato não é correto. O primeiro genoma

Teste do papel de um elemento conservado na regulação gênica

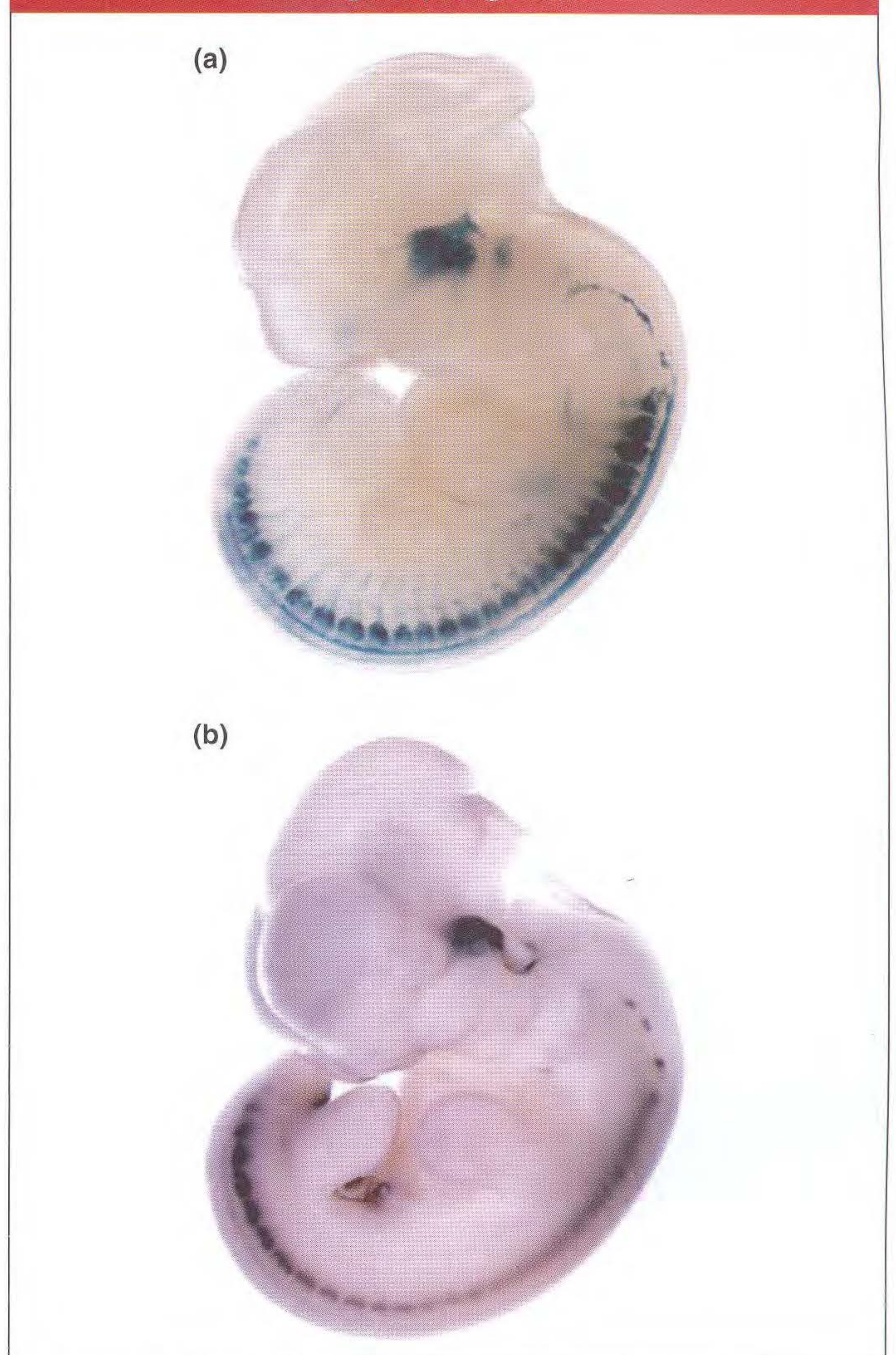


FIG. 13.16 Um elemento regulatório transcricional de ação cis é identificado em um elemento ultraconservado do genoma humano. Um elemento ultraconservado que fica perto do gene humano *ISL1* foi acoplado a um gene repórter e injetado em zigotos de camundongo. As regiões onde o gene é expresso são coradas em azul-escuro ou preto. (a) O gene repórter é expresso na cabeça e coluna dorsal de um camundongo transgênico, como visto aqui no dia 11,5 de gestação. Esse padrão de expressão corresponde ao (b) padrão nativo de expressão do gene *ISL1* de camundongo no dia 11,5 de gestação. Esse experimento demonstra como elementos não-codificantes funcionais podem ser identificados por genômica comparativa e testados em um organismo-modelo. [De G. Bejerano et al., “A Distal Enhancer and an Ultraconserved Exon Are Derived from a Novel Retroposon,” *Nature* 441, 2006, 87-90, Fig. 3.]

seqüenciado foi derivado da linhagem de *E. coli* comum de laboratório K-12. Existem muitas outras linhagens de *E. coli*, incluindo várias de importância para a saúde humana.

Em 1982, houve um surto em vários Estados de uma doença humana relacionado ao consumo de carne mal cozida. A linhagem de *E. coli* O157:H7 foi identificada como a responsável, e, desde então, ela é associada a um grande número de surtos de infecções em grande escala. De fato, existe uma estimativa de 75.000 casos por ano nos EUA. Embora

Duas linhagens de *E. coli* contêm ilhas de genes específicos de cada linhagem

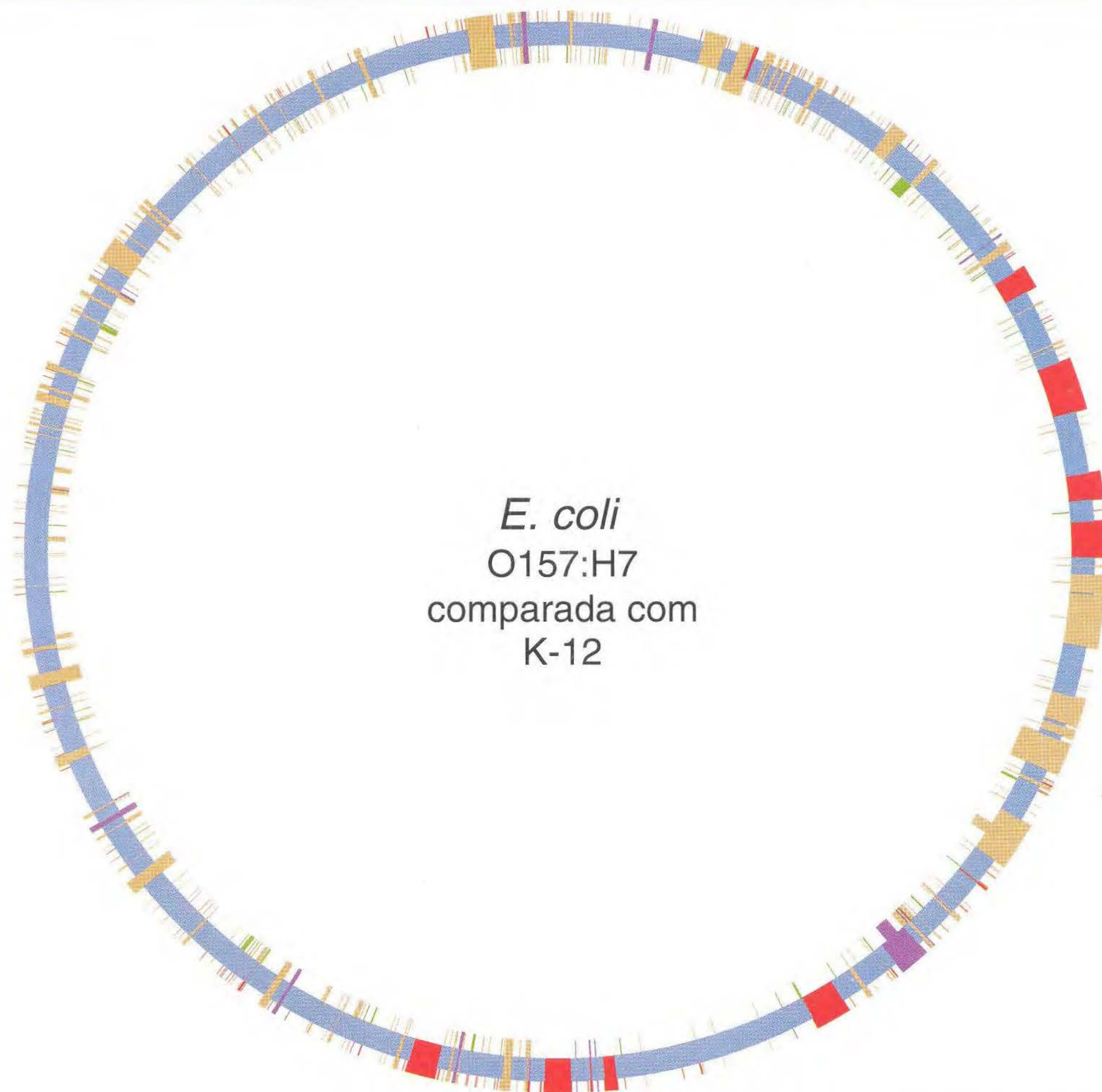


FIG. 13.17 Os mapas genômicos circulares das linhagens de *E. coli* K-12 e O157:H7. O círculo mostra a distribuição de seqüências específicas de cada linhagem. O arcabouço colinear comum a ambas as linhagens é mostrado em azul. As posições de seqüências específicas O157:H7 são mostradas em vermelho. As posições de seqüências específicas K-12 são mostradas em verde. As posições de O157:H7- e seqüências específicas K-12 no mesmo local são mostradas em laranja. As seqüências hipervariáveis são mostradas em púrpura. [De N. T. Perna et al., "Genome Sequence of Enterohaemorrhagic *Escherichia coli* O157:H7," *Nature* 409, 2001, 7529–7533. Cortesia de Guy Plunkett III e Frederick Blattner.]

a maioria das pessoas se recuperem da infecção, uma parte desenvolve a síndrome de uremia hemolítica, uma doença renal potencialmente ameaçadora à vida.

Para compreender as bases genéticas da patogenia, o genoma de uma linhagem O157:H7 de *E. coli* foi seqüenciado. As linhagens O157 e K-12 têm um arcabouço de 3.574 genes em comum codificantes da proteína, e a identidade média de nucleotídeos entre genes ortólogos é de 98,4%, comparável à de ortólogos humanos e de chimpanzés. Cerca de 25% dos ortólogos de *E. coli* codificam proteínas idênticas, similares aos 29% de ortólogos de humanos e chimpanzés.

A despeito das similaridades em muitas proteínas, os genomas e proteomas diferem enormemente em conteúdo. O genoma O157 de *E. coli* codifica 5.416 genes, enquanto o genoma de *E. coli* K-12 codifica 4.405 genes. O genoma O157 de *E. coli* contém 1.387 genes que não são encontrados no genoma de K-12, e 528 genes encontrados no genoma de K-12 não estão no genoma de O157. A comparação dos mapas genô-

micos revela que os arcabouços comuns aos dois filamentos são intercalados com ilhas de genes específicos de K-12 ou O157 (Figura 13.17).

Entre os 1.387 genes específicos de *E. coli* O157 existem muitos genes candidatos que codificam fatores de virulência, incluindo toxinas, proteínas de invasão celular, proteínas de adesão e sistemas de secreção, bem como possíveis genes metabólicos que podem ser necessários para o transporte de nutrientes, resistência a antibióticos e outras atividades que podem conferir a habilidade de sobreviver em hospedeiros diferentes. A maioria desses genes não eram conhecidos antes do seqüenciamento e não seriam conhecidos hoje se os pesquisadores fossem se basear apenas em *E. coli* K-12 como um guia para todas as *E. coli*.

O nível surpreendente de diversidade entre os membros da mesma espécie mostra como pode ser a dinâmica da evolução do genoma. A maioria dos novos genes em linhagens de *E. coli* são considerados como introduzidos por transferência

horizontal dos genomas de vírus e outras bactérias. As diferenças podem também se desenvolver por deleção gênica. Outras *E. coli* patogênicas e espécies bacterianas também tendem a exibir muitas diferenças no conteúdo gênico de suas primas não-patogênicas. A identificação dos genes que podem contribuir diretamente para a patogenicidade abriu novos caminhos para a prevenção e tratamento de doenças.

13.6 Genômica Funcional e Genética Reversa

Os geneticistas têm estudado a expressão e interações de produtos gênicos nas últimas décadas. Entretanto, esses estudos foram em pequena escala, considerando apenas um gene ou alguns genes de cada vez. Com o advento da Genômica, temos uma oportunidade de expandir esses estudos a um nível global, usando enfoques de genoma total para estudar a maioria de todos os produtos gênicos sistemática e simultaneamente. Esse enfoque global para o estudo da função, expressão e interação dos produtos gênicos é chamada de **genômica funcional**.

Ome, Sweet Ome

Além do genoma, outros conjuntos de dados globais são de interesse. Seguindo o exemplo do termo genoma, para o qual “gene” mais “oma” torna-se a palavra para “todos os genes”, os pesquisadores de genômica criaram vários termos para descrever outros conjuntos de dados globais nos quais estão trabalhando. Esse conjunto de omas inclui

O transcriptoma. Os padrões de sequência e expressão de todos os transcritos (onde, quando, quanto).

O proteoma. Os padrões de sequência e expressão de todas as proteínas (onde, quando, quanto).

O interatoma. O conjunto completo de interações físicas entre as proteínas e segmentos de DNA, entre proteínas e segmentos de RNA e entre as proteínas.

Não consideraremos todos esses omas nesta seção, mas enfocaremos algumas das técnicas globais que estão começando a ser exploradas para obter esses conjuntos de dados.

Uso de microarranjos de DNA para estudar o transcriptoma Suponha que queremos responder à pergunta: Que genes estão ativos em determinada linhagem celular sob certas condições? Essas condições podem ser um ou mais estágios no desenvolvimento, ou podem ser a presença ou ausência de um patógeno de um hormônio. Os genes ativos estão sendo transcritos em RNA, e, assim, os RNA transcritos presentes em uma célula podem nos dizer que genes estão ativos. Aqui é onde a nova tecnologia de chips de DNA usados para triar transcritos de RNA é tão poderosa.

Os chips de DNA são amostras de DNA arrumados como uma série de pontos microscópicos ligados a um “chip” de vidro do tamanho de uma lamínula de microscópio. Um chip pode conter pontos de segmentos de DNA correspondentes a todos

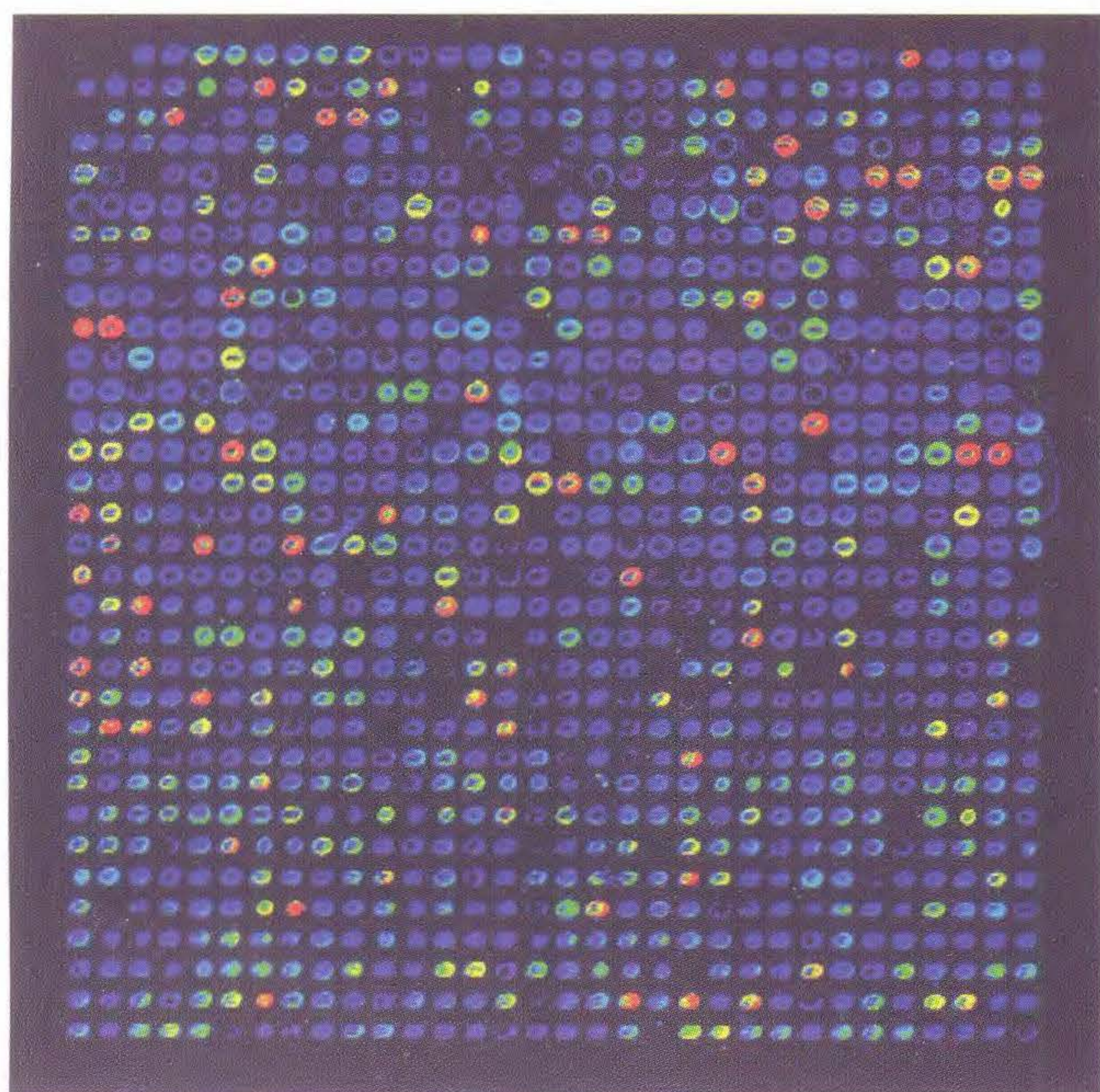
os genes em um genoma. O conjunto de DNA assim dispostos é chamado de **microarranjo**. O chip de DNA é exposto a uma amostra de RNA marcado (chamado de uma *sonda*) obtida de uma célula; cada transcrito de RNA irá ligar-se (hibridizar) à sua sequência de DNA complementar. Esses pontos no chip que estão ligados indicam quais genes estão sendo ativamente transcritos em determinada condição. Os chips de DNA revolucionaram a Genética, permitindo a triagem dos transcritos de RNA para todos os genes simultaneamente em um único experimento. Vejamos como funciona esse processo em mais detalhes.

Um protocolo para fazer chips de DNA é o seguinte. As máquinas robóticas com múltiplas pontas de impressão similares a pontas miniaturas de canetas colocam gotículas microscópicas de solução de DNA em posições específicas (endereços) no chip. O DNA é seco e tratado de modo que irá ligar-se a um vidro. Milhares de amostras podem ser aplicadas a um chip. Em um enfoque, o arranjo de DNA consiste em todos os cDNA conhecidos do genoma. Outro tipo de arranjo contém oligonucleotídeos sintéticos curtos representando a maioria ou a totalidade dos genes em um genoma. Tais arranjos são expostos a uma sonda, por exemplo, que consista no total de moléculas de RNA extraídas de um tipo particular de célula em um estágio específico no desenvolvimento. As marcações fluorescentes são ligadas à sonda, e a ligação das moléculas de sonda aos pontos homólogos de DNA no chip de vidro é monitorada automaticamente com o uso de um microscópio iluminado por um feixe de laser. Os resultados típicos são mostrados na Figura 13.18. Desse modo, os genes que estão ativos em qualquer estágio do desenvolvimento ou sob qualquer condição podem ser triados. A Figura 13.19 mostra um exemplo do perfil de desenvolvimento experimental do gene gerado pela triagem desse tipo de chip.

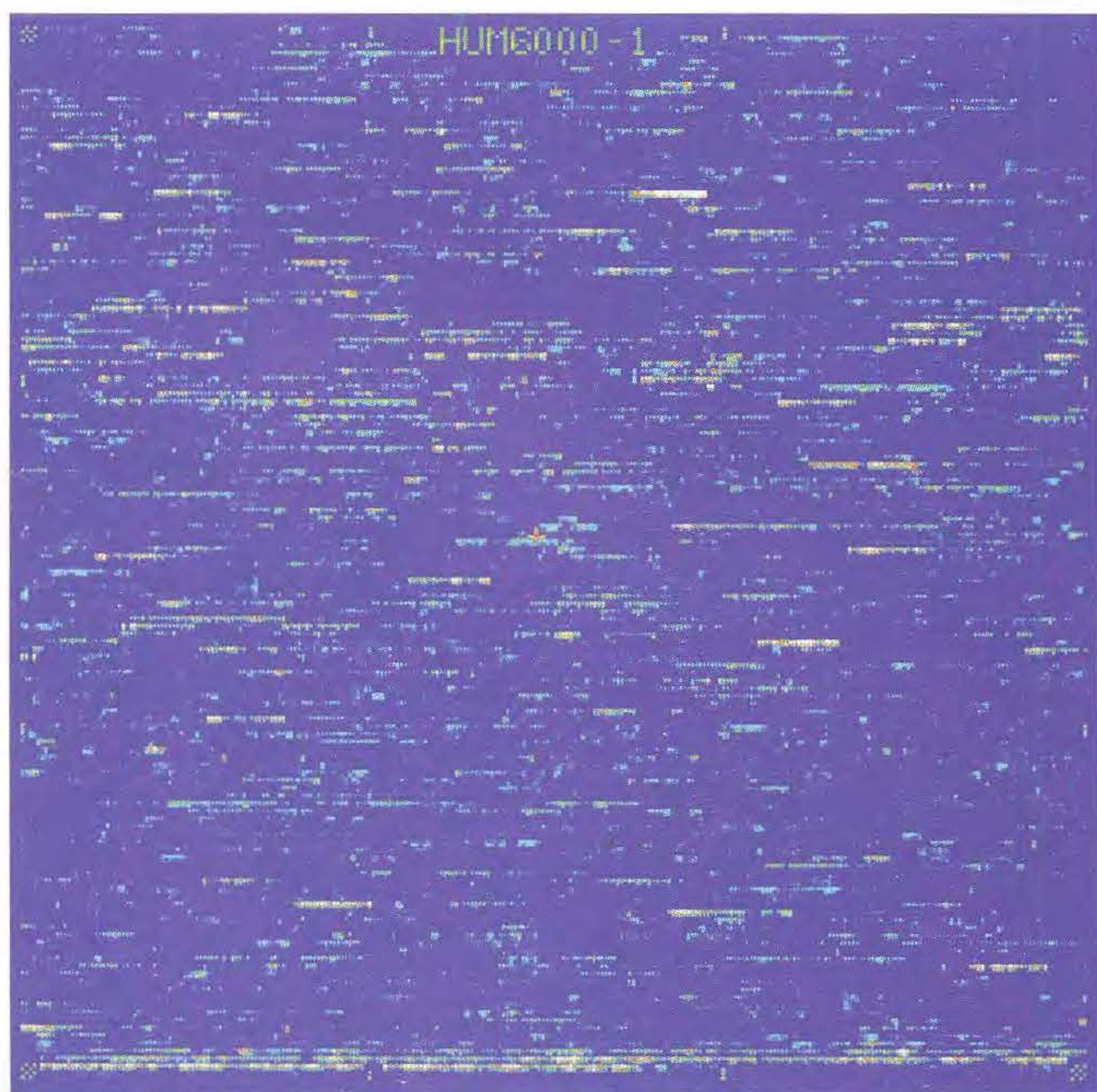
Com a compreensão de quais genes estão ativos ou inativos em determinado estágio, em determinado tipo de célula ou em várias condições ambientais, os conjuntos de genes que podem responder a sinais reguladores similares podem ser identificados. Além disso, os perfis de expressão gênica podem criar um quadro das diferenças entre células normais e aquelas com doenças. Identificando os genes cuja expressão é alterada por mutações, nas células cancerosas ou por um patógeno, os pesquisadores podem ser capazes de criar novas estratégias terapêuticas.

Usando o teste diíbrido para estudar o interatoma Uma das mais importantes atividades das proteínas é sua interação com outras proteínas. Devido ao grande número de proteínas em qualquer célula, os biólogos têm procurado modos para estudar sistematicamente todas as interações de proteínas individuais em uma célula. Um dos modos mais comuns de estudar o interatoma usa um sistema criado em células de levedura chamado de **teste diíbrido**, que detecta interações físicas entre duas proteínas. A base do teste é o ativador transcricional codificado pelo gene de levedura *GAL4* (veja o Capítulo 11). Lembre que essa proteína tem dois domínios: (1) um domínio de ligação ao DNA que se liga ao sítio de início transcricional e (2) um domínio de ativação que irá ativar a transcrição mas não pode em si ligar-se ao DNA.

O transcriptoma é estudado com o uso de microarranjos de DNA



(a)



(b)

FIG. 13.18 Detecção por fluorescência de ligação a microarranjos de DNA. As sondas são transcritas reversamente de cDNA dos mRNA. (a) Um arranjo de 1.046 cDNA com cDNA marcados com fluorescência feito de mRNA de medula óssea. O nível do sinal de hibridização segue as cores do espectro, com o vermelho sendo o mais alto e o azul o mais baixo. (b) Affymetrix GeneChip 65.000 arranjos de oligonucleotídeos representando 1.641 genes, sondados com cDNA histoespecíficos. [(a) Cortesia de Mark Scheria, Stanford University. Imagem mostrada na *Nature Genetics* 16 June 1997, p. 127, Fig. 1a. (b) Cortesia de Affymetrix Inc., Santa Clara, California. Imagem gerada por David Lockhart. Affymetrix and GeneChip são marcas registradas nos EUA usadas por Affymetrix. Imagem mostrada na *Nature Genetics* 16 June 1997, p. 127, Fig. 1b.]

Microarranjos de DNA:
usando um arranjo de
oligonucleotídeo para
analisar os padrões de
expressão gênica

W.W. ANIMAÇÃO COMPUTADORIZADA

Os microarranjos de DNA revelam perfis de expressão gênica

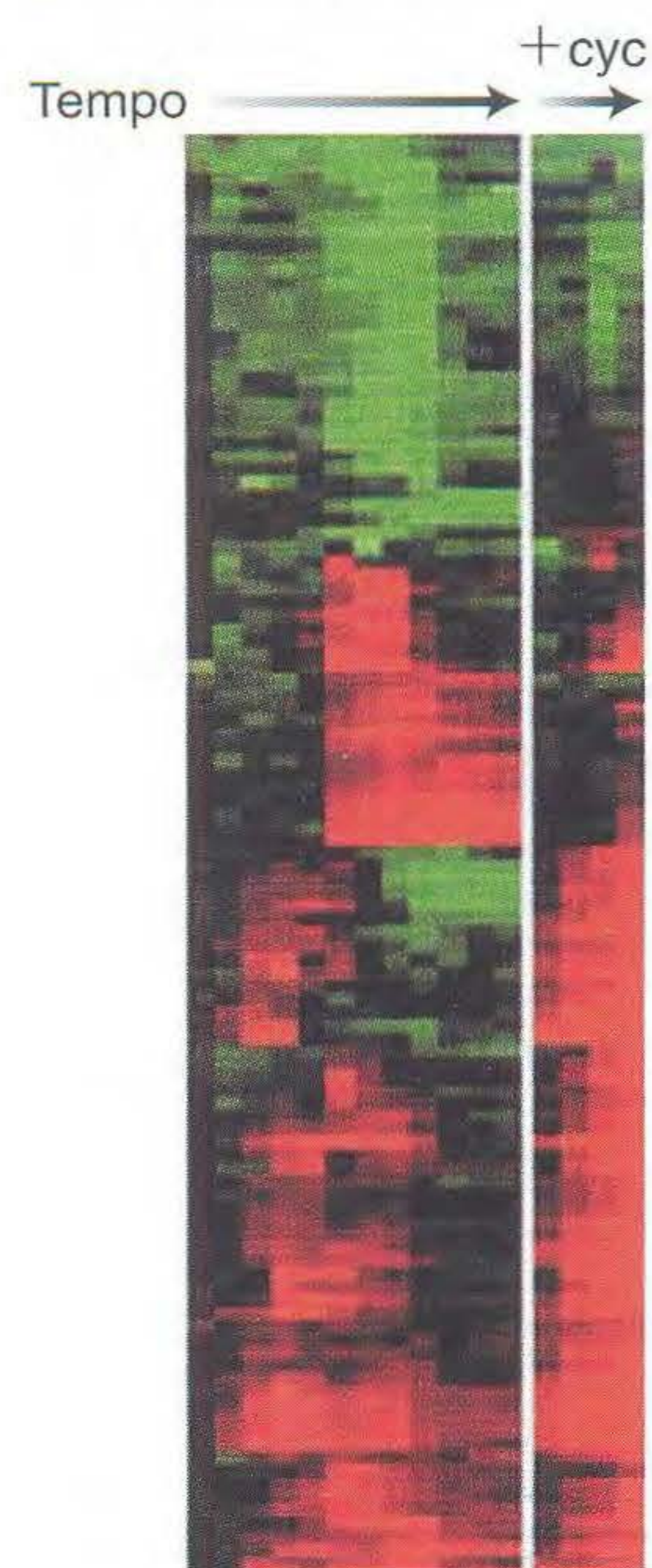


FIG. 13.19 Mostra dos padrões de expressão gênica detectados por microarranjos de DNA. Cada fileira é um gene diferente, e cada coluna é um ponto diferente. Vermelho significa que os níveis transcritos para o gene são mais altos do que no ponto inicial; verde significa que os níveis transcritos são mais baixos. As quatro colunas marcadas com + cyc são das células cultivadas com cicloeximida, significando que nenhuma síntese de proteína ocorreu nessas células. [Mike Eisen and Vishy Iyer, Stanford University. Imagem mostrada na *Nature Genetics* 18 March 1998, p. 196, Fig. 1.]

Assim, ambos os domínios devem estar em íntima proximidade para ocorrer a ativação transcricional. No sistema diíbrido, o gene para o ativador transcricional Gal4 é dividido entre dois plasmídeos de modo que um plasmídeo contém a parte codificante do domínio de ligação do DNA e o outro plasmídeo contém uma parte codificante do domínio de ativação. Em um plasmídeo, um gene para uma proteína que está sob investigação é recomposto próximo ao domínio de ligação do DNA, e essa proteína de fusão atua como uma “isca”. No outro plasmídeo, um gene para outra proteína sob investigação é recomposto para o domínio de ativação, e a proteína de fusão resultante é dita como sendo o “alvo” (Figura 13.20). Os dois plasmídeos híbridos são então introduzidos na mesma célula de levedura, talvez cruzando células haplóides contendo uma isca e plasmídeos-alvo. A etapa final é procurar pela ativação da transcrição por um conjunto de genes repórter regulado por Gal4, que seria a prova de que a isca e o alvo ligaram-se. O sistema diíbrido pode ser automatizado para possibilitar a caça das interações de proteínas ao longo do proteoma.

Estudo do interatoma usando triagem de imunoprecipitação de cromatina (ChIP) A ligação de seqüências específicas de proteínas ao DNA é crítica para a expressão correta do gene. Por exemplo, as proteínas reguladoras ligam-se aos

Estudo de interações de proteína com o uso do sistema diíbrido de levedura

Vetores diíbridos de levedura

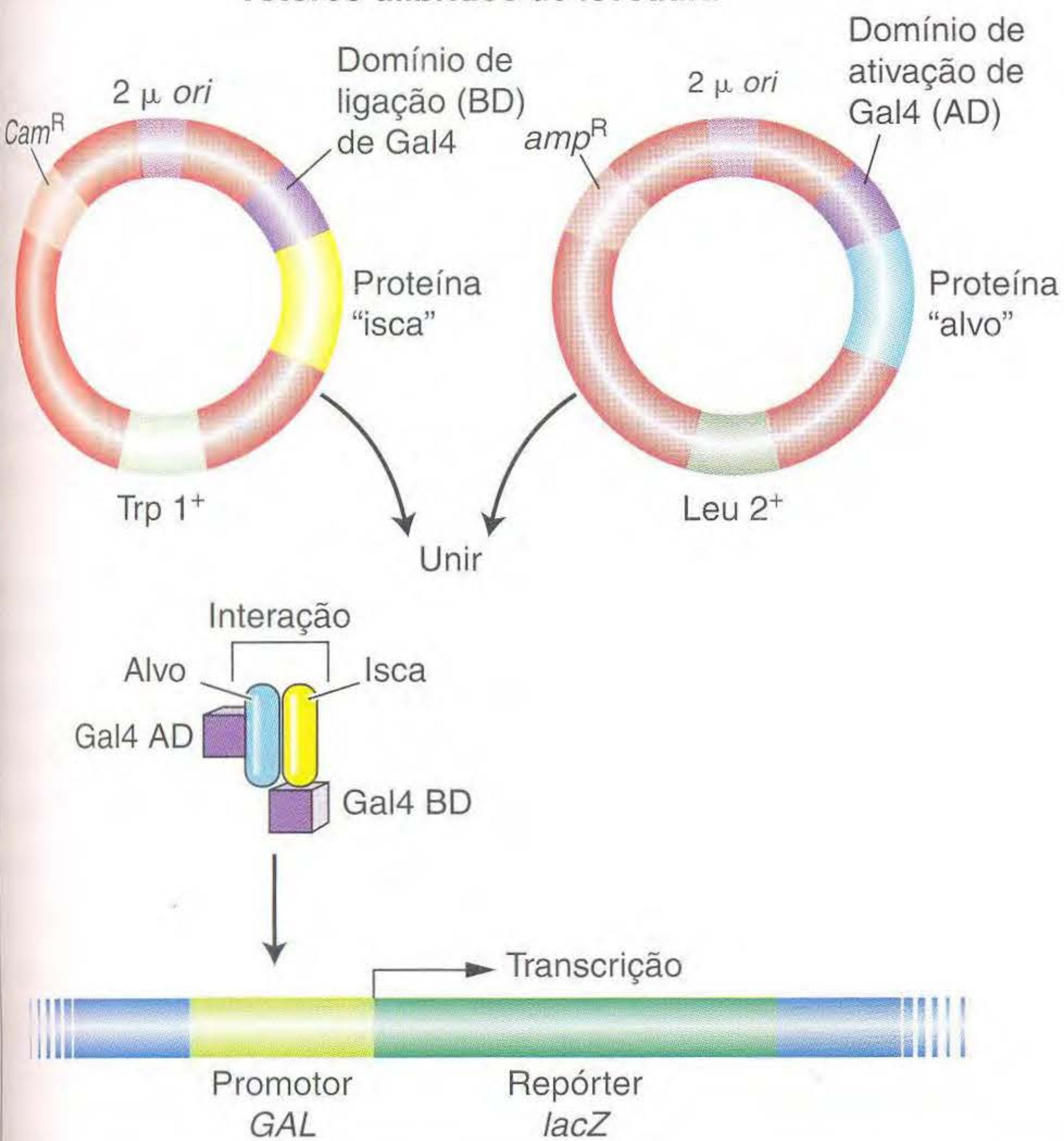


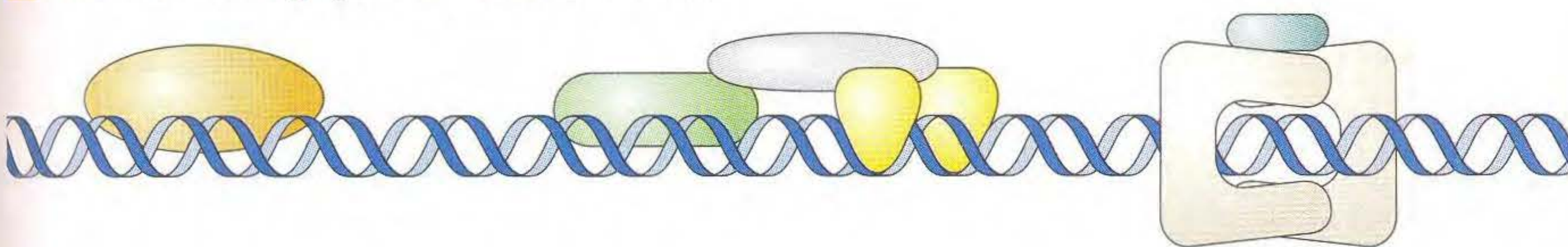
FIG. 13.20 O sistema usa a ligação de duas proteínas pesquisadas para restaurar a função da proteína Gal4, que ativa um gene repórter. Cam, Trp e Leu são componentes dos sistemas de seleção para mover o plasmídeo entre as células. O gene repórter é *lacZ*, que reside em um cromossomo de levedura (mostrado em azul).

promotores e ativam ou reprimem a transcrição tanto em bactérias quanto em eucariontes (veja os Capítulos 10, 11 e 12). No caso de eucariontes, os cromossomos são organizados em cromatina na qual a unidade fundamental, o nucleossomo, contém DNA enrolado ao redor de histonas. A modificação pós-traducional das histonas em geral determina que proteínas se ligam e onde (veja o Capítulo 11). A importância das interações proteína-DNA levou ao desenvolvimento de uma variedade de tecnologias que facilitam o isolamento de regiões específicas da cromatina, de modo que o DNA e suas proteínas associadas podem ser analisadas juntas. O método mais amplamente usado é chamado **ChIP** (para **imunoprecipitação de cromatina**), e sua aplicação é descrita a seguir (Figura 13.21).

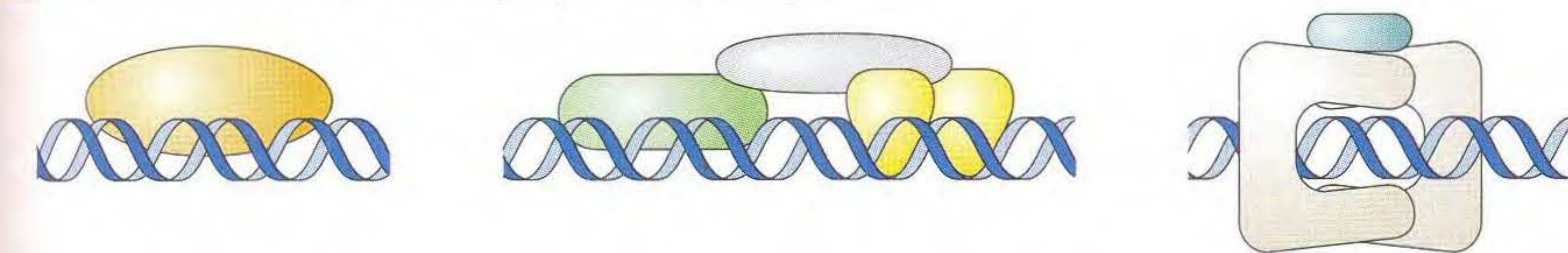
Digamos que você isolou um gene de levedura e suspeita que ele codifique uma proteína que se liga ao DNA quando a levedura está crescendo a uma alta temperatura. Você quer saber se essa proteína se liga ao DNA e, caso sim, a que sequência na levedura. Um modo para abordar essa pergunta

Etapas na análise de imunoprecipitação de cromatina (ChIP)

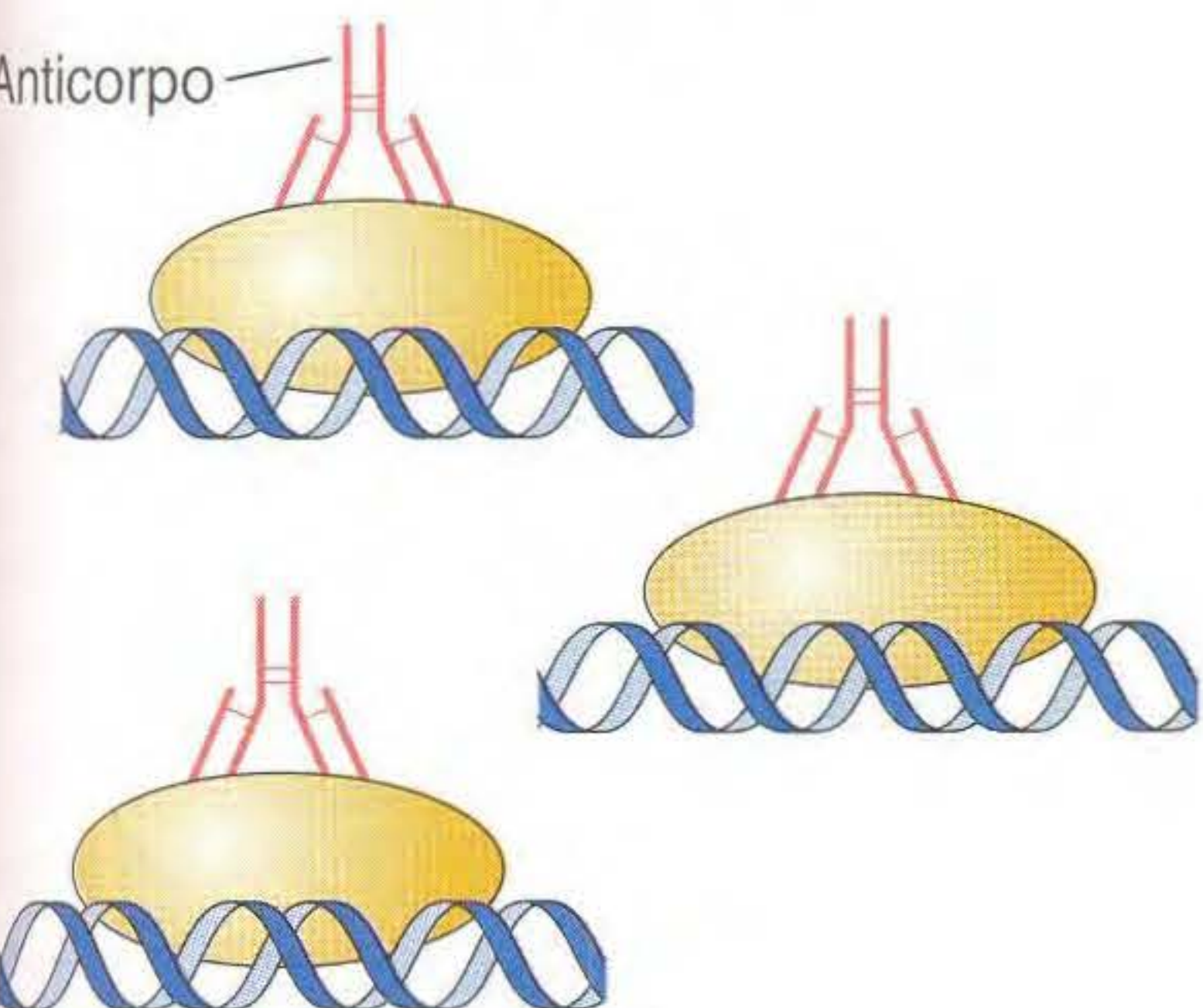
1 Proteínas de ligação cruzada ao DNA.



2 Quebra da cromatina em pequenos pedaços.



3 Adicionar anticorpo à proteína-alvo e purificar.



4 Ligação cruzada reversa para separar DNA e proteína.

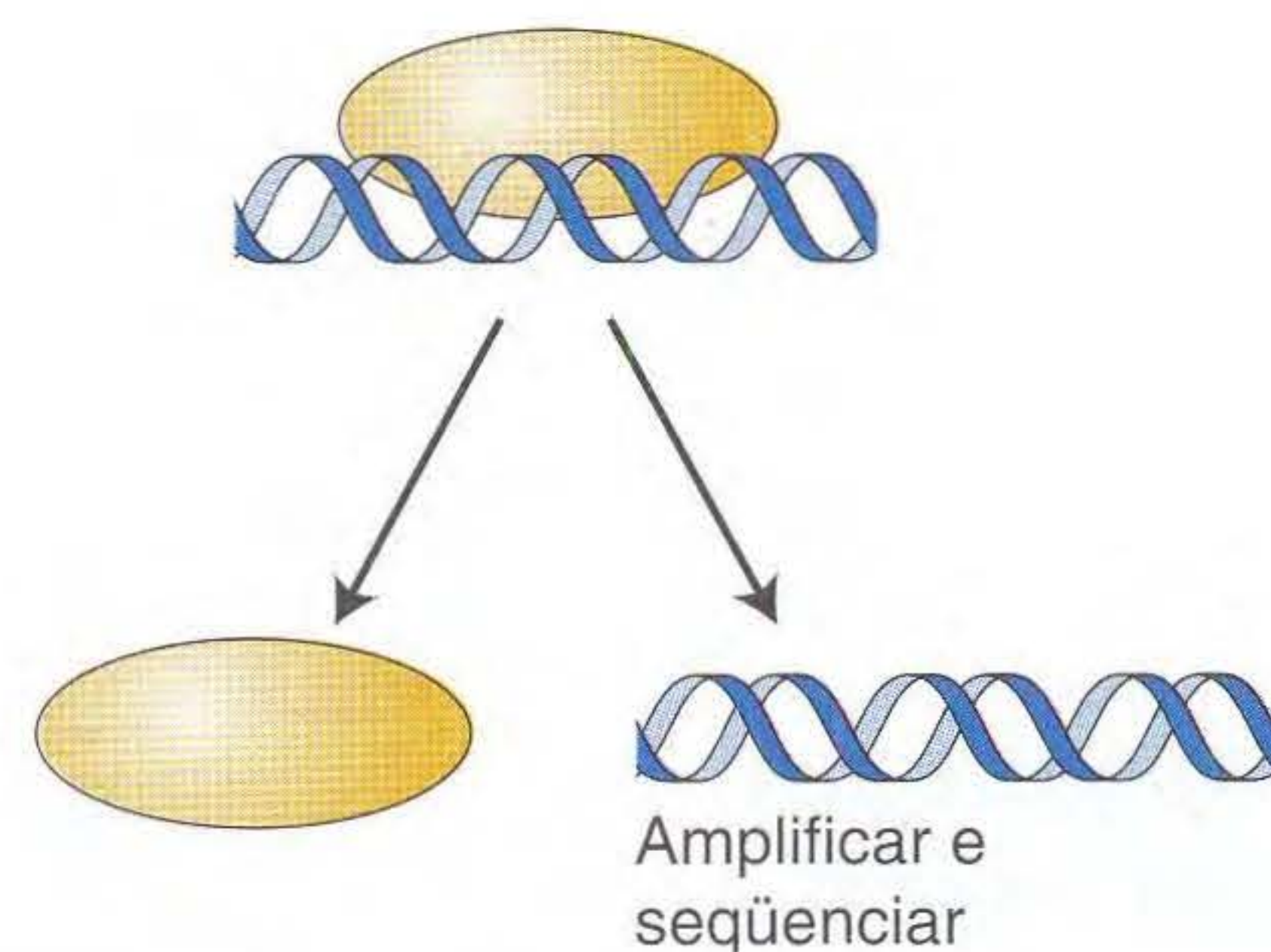


FIG. 13.21 ChIP é uma técnica para isolar o DNA e suas proteínas associadas em uma região específica de cromatina de modo que ambos possam ser analisados juntos.

é primeiro tratar as células de levedura que foram cultivadas em alta temperatura com uma substância química que faça ligação cruzada das proteínas com o DNA. Desse modo, as proteínas ligadas ao DNA na época do isolamento da cromatina permanecerão ligadas em tratamentos subsequentes. A etapa seguinte é quebrar a cromatina em pequenos pedaços. Para separar o fragmento contendo seu complexo proteína/DNA de outros, você isola anticorpos que reagem especificamente com a proteína codificada. Você adiciona seu anticorpo a uma mistura de modo que seja formado um imunoprecipitado que possa ser purificado. A proteína e o DNA podem ser analisados separadamente após a ligação cruzada ser revertida. O DNA ligado à proteína pode ser amplificado em muitas cópias, pela clonagem em bactérias ou amplificação por PCR, para preparar o sequenciamento do DNA (veja o Capítulo 20).

Como você viu no Capítulo 11, as proteínas reguladoras em geral ativam a transcrição de muitos genes, simultaneamente, ligando-se a várias regiões promotoras. Uma variação do procedimento ChIP, chamado ChIP-chip, foi criado para identificar vários sítios de ligação em um genoma sequenciado. As proteínas que se ligam a muitas regiões genômicas são imunoprecipitadas como já descrito aqui, e, após ser revertida a ligação cruzada, os fragmentos de DNA são marcados e usados para sondar chips de microarranjos que contêm, nesse caso, uma sequência genômica inteira de levedura.

O procedimento ChIP-chip também tem sido usado para decifrar o código de histonas em alguns organismos. Por exemplo, foram isolados anticorpos que reconhecem histonas com modificações pós-traducionais específicas, tais como metilação de determinada lisina (veja o Capítulo 11). O uso desse anticorpo com ChIP-chip deve identificar sequências de DNA no genoma que estão associadas a histonas que tenham essa modificação.

A Genômica e outras áreas “ômicas” geraram uma nova disciplina chamada **biologia de sistemas**. Embora o enfoque da Genética tenha sido tradicionalmente reducionista, dissecar um organismo com mutações para ver quais são as partes, a biologia de sistemas tenta juntar as partes para compreender o todo como um sistema. Um sistema biológico compreende redes de regulação gênica, cascatas de transdução de sinal, comunicação célula-célula e muitas formas de interações não só entre moléculas “genéticas”, mas com todas as outras moléculas da célula e ambiente.

Genética reversa

Os tipos de dados obtidos de experimentos de microarranjos e triagens de interação de proteínas são sugestivos de interações dentro do genoma e proteoma, mas eles não permitem que se tire uma conclusão firme sobre as funções genéticas e interações *in vivo*. Por exemplo, descobrir que a expressão de alguns genes foi perdida em alguns cânceres não é prova de causa e efeito. É necessário perturbar especificamente o funcionamento do gene e compreender os fenótipos nas condições nativas. Começando com as sequências gênicas disponíveis, os pesquisadores podem agora usar uma variedade de métodos para perturbar a função de um gene específico.

Esses métodos são chamados de **genética reversa**. A análise genética reversa começa com uma molécula conhecida, uma sequência de DNA, um mRNA ou uma proteína, e, então, tenta perturbar essa molécula para avaliar o papel do produto gênico normal na biologia do organismo.

Existem vários enfoques para a genética reversa. Um enfoque é introduzir mutações aleatórias no genoma, mas então focar o gene de interesse por identificação molecular de mutações no gene. Um segundo enfoque é conduzir uma mutagênese direcionada que produza mutações diretamente no gene de interesse. Um terceiro enfoque é criar *fenocópias*, efeitos comparáveis a fenótipos mutantes, por tratamento com agentes que interferem no mRNA ou na atividade do produto proteico final.

Cada enfoque tem suas vantagens. A mutagênese aleatória é a mais fácil de fazer, mas requer tempo e esforço para procurar em todas as mutações até encontrar a pequena proporção que inclui o gene de interesse. A mutagênese dirigida também é trabalhosa, mas, após ter sido obtida a mutação-alvo, sua caracterização é mais direta. Criar fenocópias pode ser muito eficiente, mas existem limites para os tipos de fenótipos que podem ser copiados. Consideraremos exemplos de cada um desses enfoques.

Genética reversa por mutagênese aleatória A mutagênese aleatória para genética reversa emprega os mesmos tipos de mutagênese geral que são usados para genética direta: agentes químicos, radiação ou elementos genéticos de transposição (veja o Capítulo 5). Entretanto, em vez da triagem do genoma inteiro quanto a mutações que exercem um efeito fenotípico particular, a genética reversa enfoca o gene em questão, o que pode ser feito de um dos dois modos gerais.

Um enfoque é encontrar a localização de mapa do gene. Apenas mutações que caem na região do genoma onde está situado o gene são retidas para análise molecular mais detalhada. Assim, nesse enfoque, as mutações recuperadas devem ser mapeadas. Um modo direto é cruzar um mutante novo com um mutante contendo uma deleção conhecida ou mutação do gene de interesse. Simbolicamente, o pareamento é *mutante novo/mutante conhecido*. Apenas os mutantes que resultam em prole com um fenótipo mutante (mostrando falta de complementação) são salvos para estudo.

Em outro enfoque, o gene de interesse é identificado no genoma mutagenizado e conferido quanto à presença de mutações. Por exemplo, se o mutágeno causa pequenas deleções, então, após a amplificação por PCR, os genes dos genomas parental e mutagenizado podem ser comparados, à procura de um genoma mutagenizado no qual o gene de interesse é reduzido de tamanho. Similarmente, as inserções de elementos de transposição no gene de interesse podem ser prontamente detectadas porque elas aumentam seu tamanho. As técnicas para reconhecimento de substituições de um só par também estão disponíveis. Assim, um conjunto de genomas contendo mutações aleatórias podem ser efetivamente triados para identificar a pequena fração de mutações que são de interesse para o pesquisador.

Genética reversa e mutagênese direcionada Para a maior parte do século XX, os pesquisadores encaravam a capacidade de direcionar as mutações para um gene específico como o “santo graal” inatingível da genética. Entretanto, hoje várias dessas técnicas estão disponíveis. Após um gene ter sido inativado em um indivíduo, os geneticistas podem avaliar o fenótipo exibido como indício da função do gene. De um modo geral, as ferramentas para as mutações gênicas dirigidas são baseadas em técnicas genéticas desenvolvidas para organismos-modelo. Assim, embora a perturbação dos genes de leveduras, mosca ou camundongo de um modo direto e eficiente seja factível, tal perturbação não pode ser feita em muitas espécies não-modelo.

A mutagênese específica de genes em geral requer a substituição de uma cópia residente tipo selvagem de todo um gene por uma versão mutada desse gene. O gene mutado é inserido no cromossomo por um mecanismo similar à recombinação homóloga, substituindo a sequência normal pelo mutante (Figura 13.22). Esse enfoque pode ser usado para o nocaute gênico direcionado, no qual um alelo nulo substitui a cópia tipo selvagem. Algumas técnicas são tão eficientes que, em *E. coli*, por exemplo, estão sendo feitos esforços para mutar sistematicamente cada gene no genoma de K-12 para avaliar sua função biológica.

A mutagênese direcionada é o modo mais preciso para obter mutações em um gene específico, e agora pode ser praticada em uma variedade de sistemas-modelo, incluindo camundongos e moscas.

Genética reversa por fenocópia A vantagem de inativar um gene é que as mutações serão transmitidas de uma geração para a seguinte, e, assim, uma vez obtida uma linhagem de mutantes, ela está sempre disponível para estudos futuros. Entretanto, apenas organismos bem desenvolvidos como modelos moleculares podem ser usados para tais manipulações. Por outro lado, as fenocópias podem ser aplicadas a mui-

Perturbar o funcionamento gênico com o uso de mutagênese direcionada

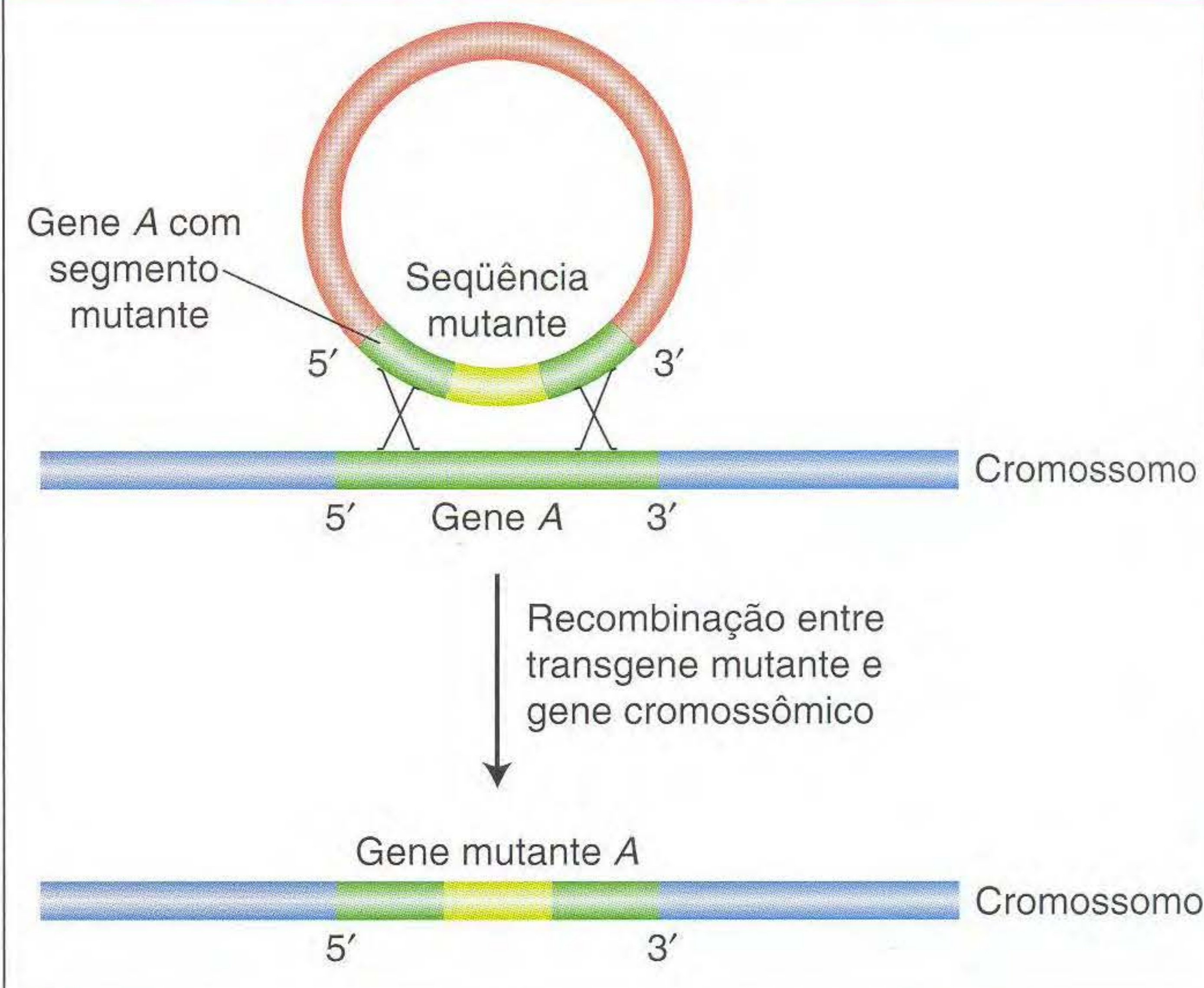


FIG. 13.22 O evento molecular básico na substituição gênica direcionada. Um transgene contendo seqüências de duas pontas de um gene, mas com um segmento selecionável de DNA entre elas é introduzido em uma célula. A recombinação dupla entre o transgene e um gene cromossômico normal produz um gene recombinante cromossômico que incorporou o segmento anormal.

tos organismos independentemente de quanto a tecnologia genética está desenvolvida para determinada espécie. As duas seções seguintes descrevem duas técnicas de fenocópia.

RNA de interferência Um achado importante nos últimos 10 anos foi a descoberta de um mecanismo generalizado cuja função natural parece ser proteger a célula de um DNA exógeno. Esse mecanismo é chamado de **RNA de interferência (RNAi)**, descrito no Capítulo 8. Os pesquisadores se aprovei-

Perturbação da função gênica com o uso de RNA de interferência

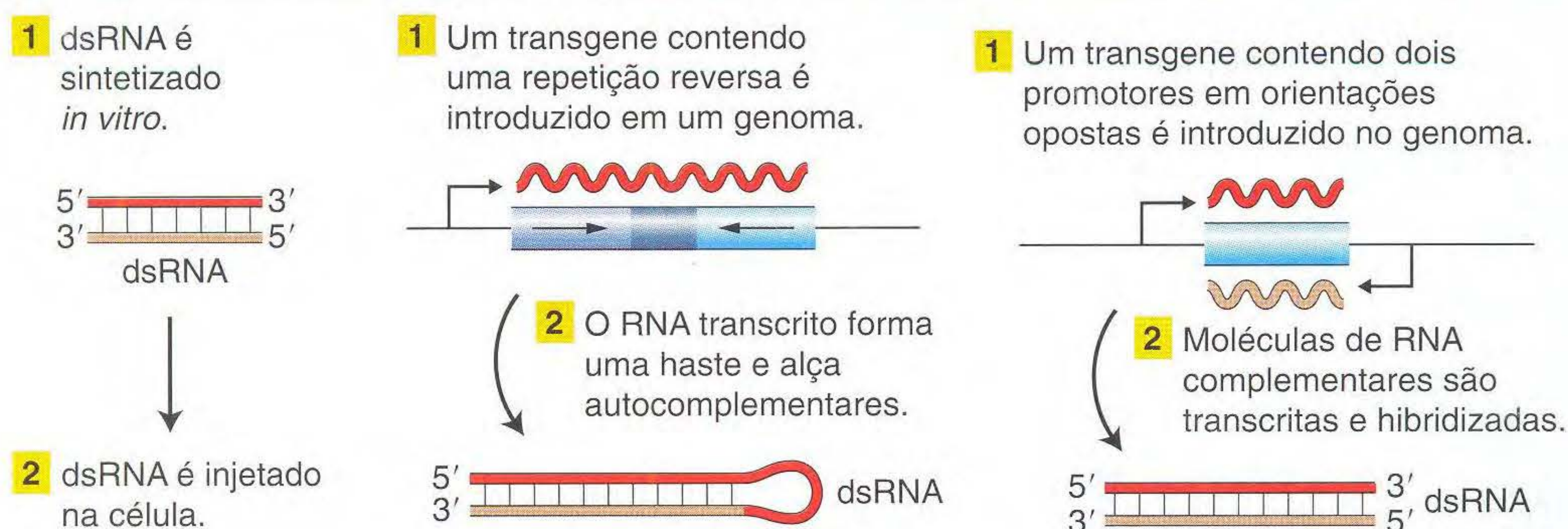


FIG. 13.23 Três modos para criar e introduzir RNA bifilamentar (dsRNA) em uma célula. O dsRNA irá então estimular o RNAi, degradando seqüências que correspondem às do dsRNA. [Reimpresso com permissão de S. Hammond, A. Caudy, and G. Hannon, *Nat. Rev. Genet.* 2, 2001, 116.]

taram desse mecanismo celular para fazer um método poderoso para inativar genes específicos. A inativação é obtida do seguinte modo. Um RNA bifilamentar é feito com seqüências homólogas à parte do gene em estudo e é introduzido na célula (Figura 13.23). O complexo de silenciamento induzido pelo RNA, ou RISC, então degrada qualquer mRNA nativo que é complementar ao RNA bifilamentar. O resultado final é uma redução considerável dos níveis de mRNA que dura horas ou dias, nulificando assim a expressão desse gene. A técnica foi aplicada bem-sucedidamente em vários sistemas-modelo, incluindo *C. elegans*, *Drosophila*, peixe-zebra e várias espécies de plantas.

Para aplicar as técnicas de fenocópias a organismos não-modelo, os genes-alvo podem ser identificados pela genômica comparativa. Então, as seqüências de RNAi são produzidas para ter como alvo a inibição de genes-alvo específicos. Essa técnica já foi aplicada a um mosquito que leva malária (*Anopheles gambiae*). Usando essas técnicas, os cientistas podem compreender melhor os mecanismos biológicos relativos ao efeito médico ou econômico de tal espécie. Por exemplo, os genes que controlam o ciclo de vida complicado do parasita da malária, parcialmente dentro do mosquito hospedeiro e parcialmente dentro do corpo humano, podem ser mais bem compreendidos, revelando novos modos de controle da doença infecciosa mais comum do mundo.

Genética química Um outro estágio do processo de transferência de informação que pode ser direcionado para fenocópia é a própria proteína. Foi desenvolvida uma técnica em escala genômica para esse fim, que recebeu o nome de **genética química**. Essa técnica, amplamente usada na indústria farmacêutica, é baseada em reduzir a atividade de uma proteína produzida por um gene-alvo pela ligação de uma pequena molécula inibitória (Figura 13.24). Com o uso da robótica, bibliotecas de milhares de pequenas moléculas sintéticas relacionadas podem ser testadas quanto a sua habilidade em se ligar fortemente a uma proteína específica e, assim, inibir sua atividade *in vitro*. Uma molécula promissora pode então ser introduzida em células e testada quanto à sua habilidade de alterar a função. Se um composto inibe suficientemente a atividade da proteína, então uma célula ou um organismo podem ser tratados com esse composto químico para obter uma fenocópia do fenótipo mutante para o gene-alvo.

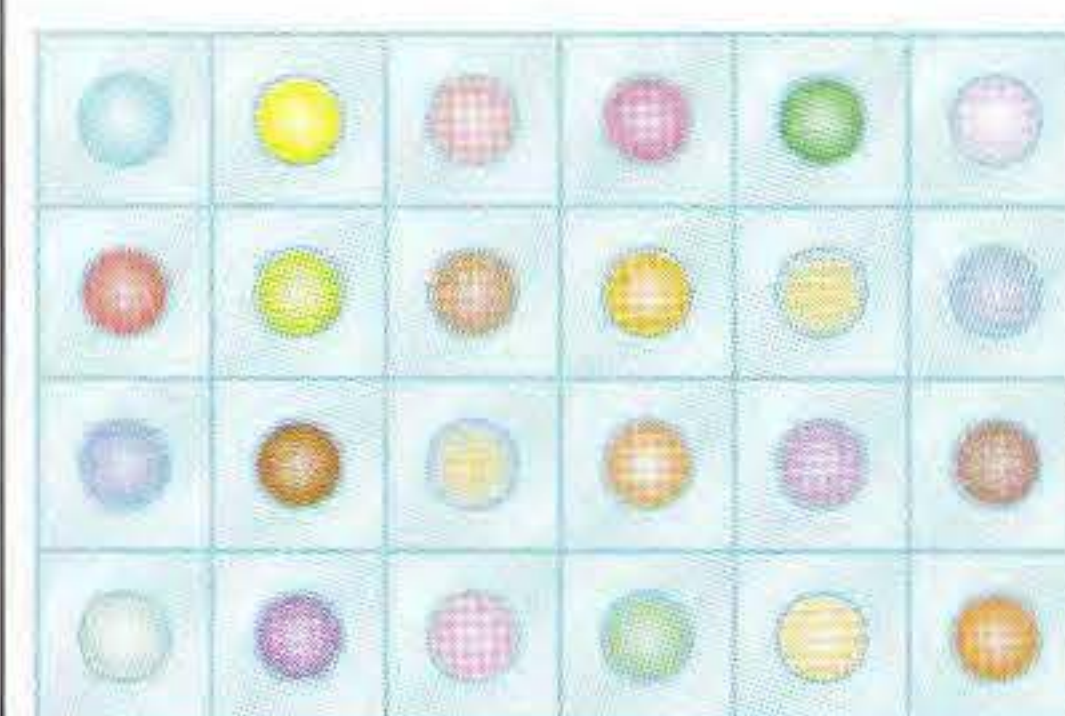
A despeito de seu nome, a genética química não é uma técnica genética, pois não envolve herança. É uma extensão sistemática do uso de longa duração de drogas inibidoras (uma forma de fenocopiar) para inativar uma proteína em um processo bioquímico específico na célula. O problema com a maioria das drogas inibidoras é que elas não são 100% específicas de uma única proteína, e, assim, inadvertidamente, elas inibem múltiplas proteínas e múltiplos processos bioquímicos em um organismo, causando ambiguidades que dificultam interpretar os resultados. Pelo uso de bibliotecas químicas e testes de robótica quanto à especificidade, a genômica química promete desenvolver compostos com muito mais especificidade, eficácia e segurança que os métodos tradicionais de triagem de drogas.

Perturbação da função proteica com o uso de genética química

(a) Genética química direta



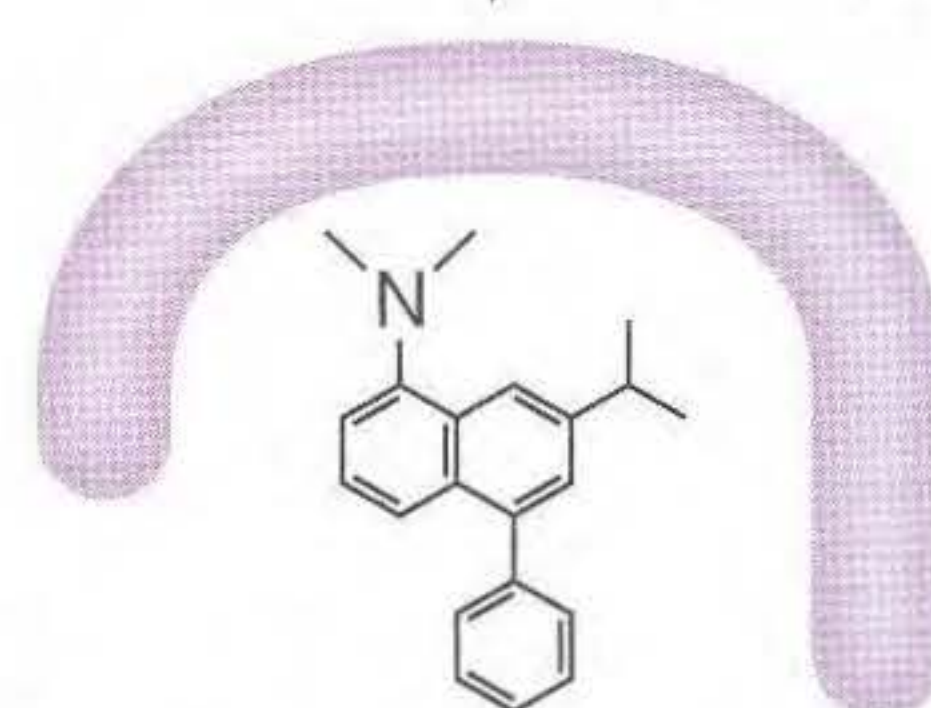
Adicionar um composto por poço.



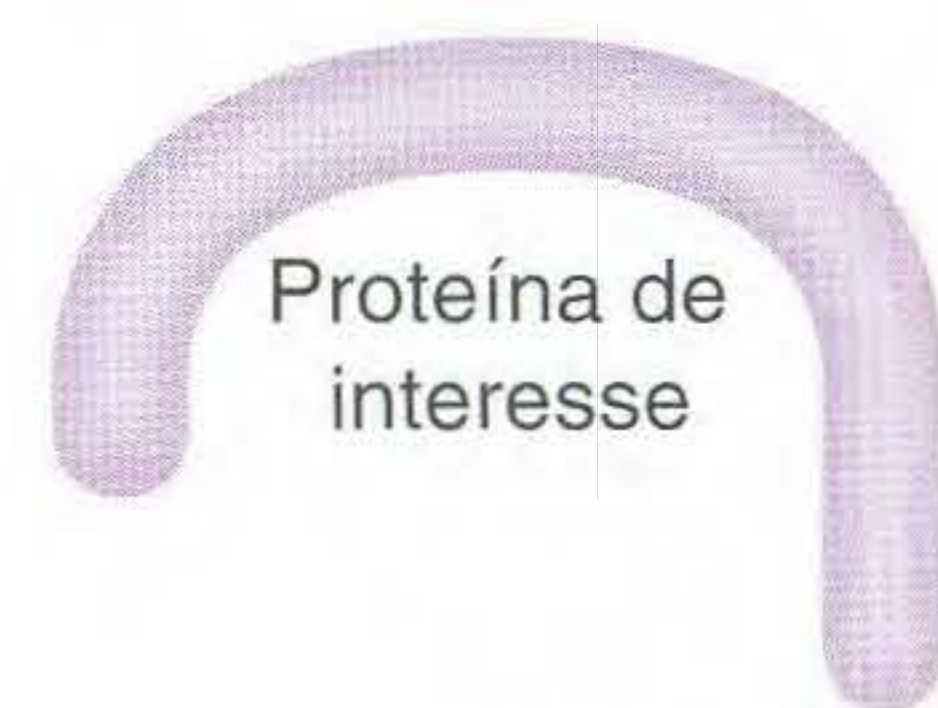
Encontrar composto que produz fenótipo de interesse.



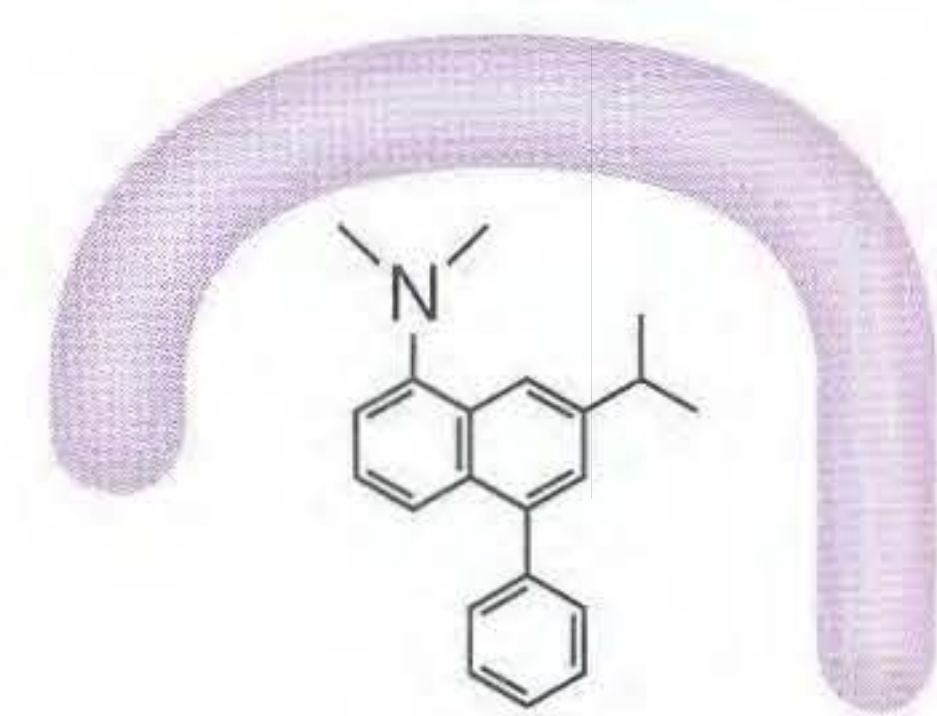
Identificar a proteína-alvo do composto.



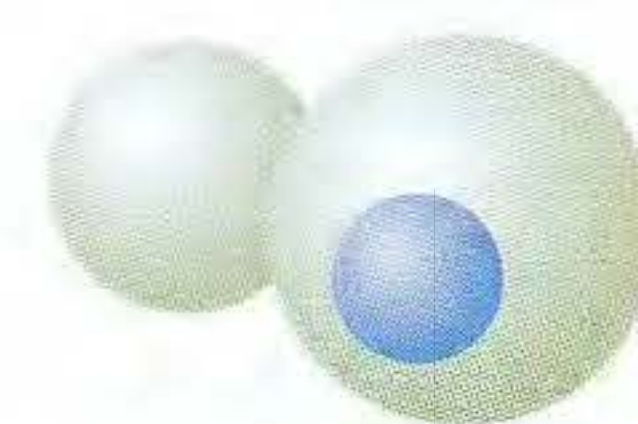
(b) Genética química reversa



Triagem de compostos que estão ligados a proteína.



Tratar células com molécula que se liga a proteína.



Analisar quanto a fenótipo.

FIG. 13.24 A genética química é usada para reduzir a atividade da proteína-alvo de um gene pela ligação de uma pequena molécula inibidora. (a) Um exemplo de genética química *direta*, no qual pequenas moléculas são diretamente testadas em células de leveduras para identificar uma que produza um fenótipo de interesse. (b) Um exemplo de genética química *reversa*, no qual uma pequena molécula primeiro é mostrada ligando-se a uma proteína de interesse, sendo subsequentemente testada por seu efeito fenotípico quando aplicada às células. [De B. Stockwell, "Chemical Genetics: Ligand-Based Discovery of Gene Function," *Nat. Rev. Genet.*, 2000, 117.]

O RNAi e a genômica química fornecem meios de interferir experimentalmente na função de um gene específico, sem mudar sua seqüência de DNA (geralmente chamada de *fenocopiar*).

Genômica funcional com organismos não-modelo Grande parte de nossa consideração da dissecação mutacional e fenocópia

Inserção de transgenes em organismo não-modelo

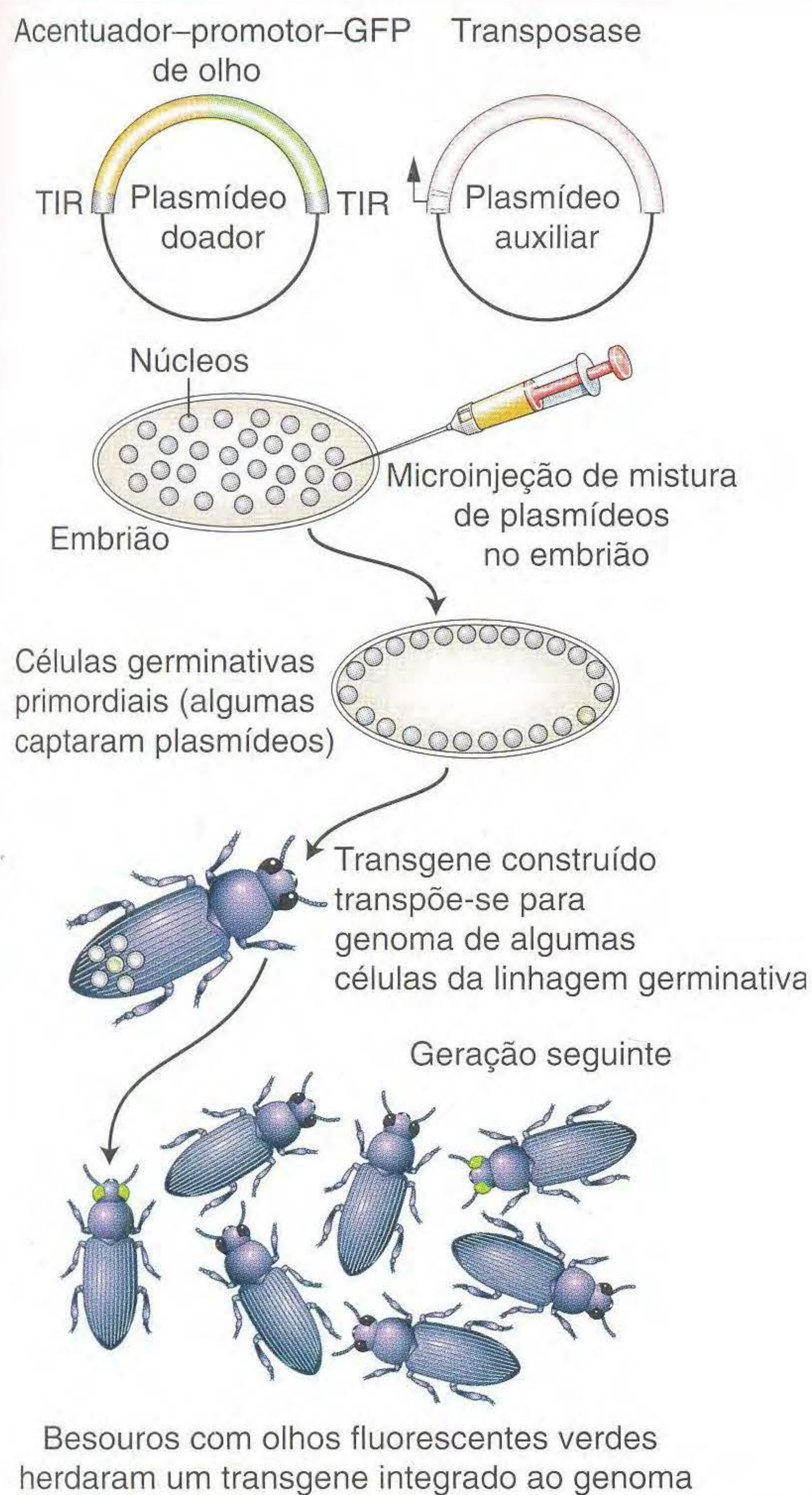


FIG. 13.25 Criação de besouros transgênicos expressando uma proteína verde fluorescente. TIR, repetição terminal invertida. [De E. A. Wimmer, "Applications of Insect Transgenesis," *Nat. Rev. Genet.* 4, 2003, 225-232.]

piar enfocou organismos-modelo genéticos. Um dos próximos desafios é aplicar esses sistemas mais amplamente, incluindo espécies que têm efeitos negativos na sociedade humana, tais como parasitas, portadores de doenças ou pestes agrícolas. As técnicas de Genética clássica não são prontamente aplicáveis à maioria dessas espécies, mas os papéis de genes específicos podem ser avaliados por transgênese ou fenocopiagem.

O primeiro enfoque, a inserção de transgenes, é mostrado na Figura 13.25. Esse exemplo envolve besouros, muitos dos quais são pestes agrícolas. Nesse caso, os transgenes foram inseridos aleatoriamente no genoma de besouro. Os besouros transgênicos podem ser produzidos usando metodologia similar à usada para produzir drosófilas transgênicas (veja o Capítulo 20). Entretanto, é necessário algum modo para identificar a transgênese bem-sucedida. Portanto, a técnica depende do uso de um gene repórter que possa ser expresso em um receptor tipo selvagem. A proteína fluorescente verde (GFP), originalmente isolada de água-viva, é um útil repórter para essa aplicação. Como em *Drosophila*, os transgenes são inseridos como partes dos transposons, e um plasmídeo auxiliar codificando a transposase facilita a inserção do transposon portador do transgene. A Figura 13.25 mostra o uso de transgenes GFP ativados por um elemento acentuador que estimula a expressão no olho do inseto. Esse método também foi efetivamente usado para criar transgenes que expressam GFP na espécie de mosquito que leva à febre amarela e à febre da dengue (*Aedes aegypti*), um besouro da farinha (*Tribolium castaneum*) e o bicho-da-seda (*Bombyx mori*) (Figura 13.26).

Exemplos de insetos não-modelo que expressam um transgene

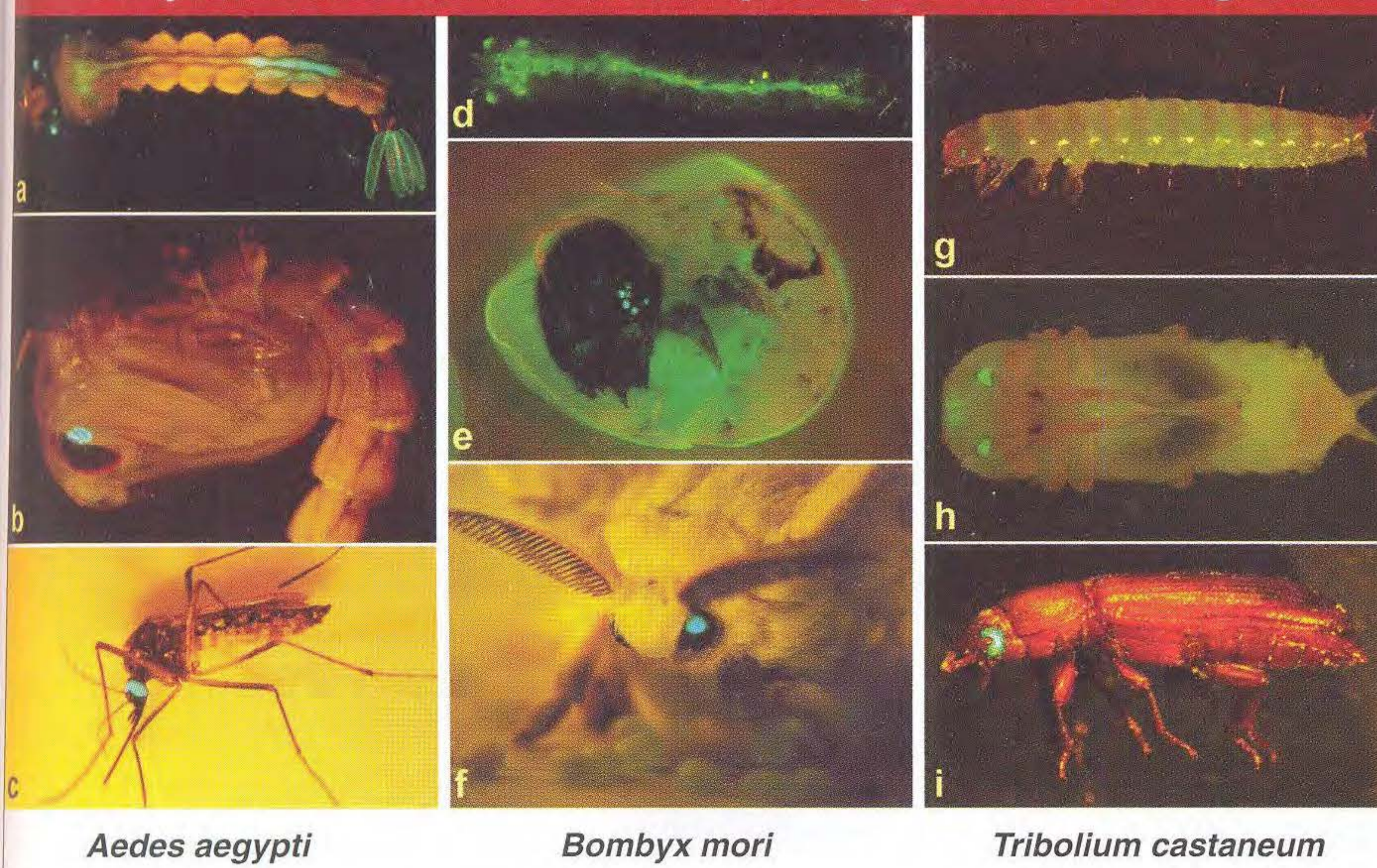


FIG. 13.26 Exemplos de uma proteína repórter verde fluorescente transgênica expressa nos olhos de alguns insetos não-modelo. A expressão é ativada a partir de um único promotor ativo no olho. Os insetos são o mosquito (*Aedes aegypti*), bicho-da-seda (*Bombyx mori*) e besouro (*Tribolium castaneum*). [(a-c) Cortesia de V. A. Kooks and Alexander S. Raikhel. (d) De J. L. Thomas et al. Copyright 2002 by Elsevier Science. (e e f) Cortesia de Marek Jindra. (g-i) Copyright 2000 by Elsevier Science.]

► Resumo

A análise genômica usa os enfoques da análise genética e os amplia para a coleta dos conjuntos de dados globais para atingir as metas tais como o mapeamento e seqüenciamento de todo o genoma e a caracterização de todos os transcritos e proteínas. As técnicas genômicas requerem o rápido processamento de grandes quantidades de material experimental, todos dependentes de ampla automação.

O problema principal em compilar uma seqüência precisa de um genoma é relacionar curtas seqüências de leitura uma à outra por identidade de seqüência para acumular uma seqüência de consenso de um genoma inteiro. Isso pode ser feito de modo direto para genomas bacterianos ou de archaea alinhando seqüências superpostas de leituras diferentes de seqüências para compilar todo o genoma, pois existem poucos segmentos de DNA, ou nenhum, presentes em mais de uma cópia em procariontes. Entretanto, genomas complexos estão repletos de tais seqüências repetitivas. Essas seqüências repetitivas interferem na produção de seqüências contig exatas. O problema é resolvido ou pelo seqüenciamento shotgun de todo o genoma, com o uso de leituras de extremidades pareadas, ou pelo seqüenciamento de clones ordenados, que trata elementos repetitivos dispersos como únicos no contexto de um clone. Ao contrário do seqüenciamento WGS, o seqüenciamento clone a clone requer que seja produzido o mapa físico da distribuição de clones ordenados e orientados.

Ter o mapa da seqüência genômica fornece a matéria-prima, o texto criptografado do genoma. A tarefa da bio-

informática é interpretar essa informação criptografada. Para a análise dos produtos gênicos, são usadas técnicas computacionais para identificar matrizes de leitura abertas e RNA não-codificantes, e, então, integrar esses resultados com evidências experimentais disponíveis para transcrever estruturas (seqüências de cDNA), similaridades de proteínas e conhecimento de motivos de seqüências características.

Um dos meios mais poderosos para avançar a análise e anotação dos genomas é pela comparação com os genomas de espécies correlatas. A conservação de seqüências entre espécies é um guia confiável para identificar seqüências funcionais nos genomas complexos de muitos animais e plantas. A genômica comparativa também pode revelar como os genomas mudaram no curso da evolução e como essas mudanças podem se relacionar a diferenças em fisiologia, anatomia ou comportamento entre as espécies. Na genômica bacteriana, as comparações de linhagens patogênicas e não-patogênicas revelaram muitas diferenças no conteúdo gênico que podem contribuir para a patogenicidade.

A genômica funcional tenta compreender o funcionamento do genoma como um sistema integral. Dois elementos principais são o transcriptoma, o conjunto de todos os transcritos produzidos, e o interactoma, o conjunto de produtos gênicos que interagem e outras moléculas que, juntas, possibilitam que a célula seja produzida e funcione. A função dos genomas individuais e os produtos gênicos para os quais as mutações clássicas não estão disponíveis podem ser testados por genética reversa – por mutação direcionada ou fenocopiagem.

► Termos-chave

anotação
arcabouço
biblioteca genômica
bioinformática
biologia de sistemas
ChIP (análise de imunoprecipitação de cromatina)
clone contig
genética química
genética reversa
genômica

genômica comparativa
genômica funcional
homólogo
leituras de pontas pareadas
mapa físico
matriz de leitura aberta (ORF)
microarranjo
montagem de seqüência
ortólogo
parálogo
projeto genoma

proteoma
pseudogene
pseudogene processado
RNA de interferência (RNAi)
seqüência contig
seqüência de consenso
seqüência expressa marcada (EST)
sintenia
supercontig
teste diíbrido
vetor

► Problemas Resolvidos

Problema resolvido 1. Você quer estudar o desenvolvimento de um sistema olfativo (receptor de cheiro) no camundongo. Você sabe que as células que sentem odores químicos específicos (odorantes) estão situadas no revestimento das passagens nasais do camundongo. Descreva alguns enfoques para o uso da genética reversa para estudar a olfação.

SOLUÇÃO

Existem muitos enfoques que podem ser imaginados. Para a genética reversa, você poderia querer identificar genes candidatos que são expressos nos revestimentos das passagens nasais. Considerando as técnicas da genômica funcional, essa identificação poderia ser feita purificando o RNA de células isoladas das passagens nasais e usando esse RNA como uma

sonda de chips de DNA contendo seqüências que correspondem a todos os mRNA no camundongo. Por exemplo, você pode escolher primeiro examinar os mRNA que são expressos no revestimento da passagem nasal, mas em nenhuma outra parte no camundongo, como importantes candidatos para um papel específico na olfação. (Muitas das moléculas importantes também podem ter outros papéis em outras partes do corpo, mas você tem que começar em alguma parte.)

Alternativamente, você pode escolher começar com genes cujos produtos proteicos são proteínas candidatas a ligar os próprios odorantes. Independentemente de sua escolha, a etapa seguinte seria construir um nocaute direcionado do gene que codifica cada mRNA ou proteína de interesse, ou usar RNA anti-sentido ou bifilamentar injetado para tentar uma fenocópia do fenótipo de perda de função de cada um dos genes candidatos.

Problemas

PROBLEMAS BÁSICOS

1. A palavra *contig* é derivada da palavra *contíguo*. Explique a derivação.
2. Explique o enfoque que você aplicaria ao seqüenciamento do genoma de uma espécie de bactérias recém-descoberta.
3. As leituras de seqüenciamento terminal das inserções de clones são uma parte rotineira do seqüenciamento do genoma. Como é obtida a parte central do clone inserido?
4. Qual a diferença entre um contig e um arcabouço?
5. Dois determinados contigs são suspeitos de estar adjacentes, possivelmente separados por DNA repetitivo. Em uma tentativa de ligá-los, são usadas as seqüências finais como primers para tentar cobrir o espaço. Esse enfoque é razoável? Em que situação ele não funciona?
6. Um segmento de DNA clonado contendo um gene codificante de proteína é marcado radioativamente e usado na hibridização *in situ*. A radioatividade foi observada em cinco regiões em cromossomos diferentes. Como é possível esse resultado?
7. Em um experimento de hibridização *in situ*, determinado clone está ligado apenas ao cromossomo X em um menino sem sintomas de doença. Entretanto, em um menino com distrofia muscular de Duchenne (doença recessiva ligada ao X), ele se liga a um cromossomo X e a um autossomo. Explique. Esse clone pode ser útil em isolar o gene para a distrofia muscular de Duchenne?
8. Em uma análise genômica procurando um gene específico, foi encontrado um gene candidato tendo uma única substituição de par de bases resultando em uma mudança de aminoácidos não-sinônimos. O que você teria que verificar antes de comemorar?
9. O operador de bactérias é um sítio de ligação?
10. Determinado cDNA com 2 kb de tamanho é hibridizado a oito fragmentos genômicos do tamanho total de 30 kb e contém dois EST curtos. Os EST também foram encontrados em dois dos fragmentos genômicos, cada um com 2 kb. Esquematize uma explicação possível para esses resultados.
11. Um fragmento seqüenciado de DNA de *Drosophila* foi usado em uma pesquisa BLAST. A melhor correspondência (mais próxima) foi com um gene de cinase de *Neurospora*. Essa correspondência significa que a seqüência de *Drosophila* contém um gene de cinase?
12. Em um teste diíbrido, determinado gene A deu resultados positivos com dois clones, M e N. Quando M foi usado, ele deu positivos com três clones, A, S e Q. O clone N deu apenas um positivo (com A). Desenvolva uma interpretação hipotética desses resultados.
13. Você tem as seguintes seqüências de um clone genômico do genoma de *Drosophila melanogaster*:
 Seqüência 1: TGGCCGTGATGGGCAGTTCCGGTG
 Seqüência 2: TTCCGGTGCCGGAAGA
 Seqüência 3: CTATCCGGGCGAACTTTTGGCCG
 Seqüência 4: CGTGATGGGCAGTTCCGGTG
 Seqüência 5: TTGGCCGTGATGGGCAGTT
 Seqüência 6: CGAACTTTTGGCCGTGATGGGCAGTTCC
 Use essas seis seqüências para criar uma seqüência contig dessa parte do genoma de *D. melanogaster*.
14. Às vezes, os cDNA viram “monstros”; isto é, as fusões de cópias de DNA de dois mRNA diferentes acidentalmente inseridos adjacentes um ao outro no mesmo clone. Você suspeita que um clone de cDNA do nemátodo *Caenorhabditis elegans* é um dos monstros porque a seqüência de cDNA inserida prevê uma proteína com dois domínios estruturais não normalmente observados na mesma proteína. Como você usaria a habilidade de toda a seqüência genômica para avaliar se esse clone de cDNA é um monstro ou não?
15. Você seqüenciou o genoma das bactérias *Salmonella typhimurium*, e usou a análise BLAST para identificar similaridades dentro do genoma de *S. typhimurium* para proteínas conhecidas. Você encontrou uma proteína que é 100% idêntica à bactéria *Escherichia coli*. Quando você compara as seqüências de nucleotídeos de *S. typhimurium* e genes de *E. coli*, você descobre que suas seqüências de nucleotídeos são apenas 87% idênticas.
 a. Explique essa observação.
 b. O que essas observações lhe dizem sobre os méritos das pesquisas de similaridade de proteínas *versus* nucleotídeos na identificação de genes correlatos?
16. Para inativar um gene por RNAi, que informação você precisa? Você precisa da posição de mapa do gene-alvo?
17. Descreva dois métodos diferentes usados para gerar fenocópias. Qual a finalidade de gerar fenocópias?
18. Qual a diferença entre genéticas direta e reversa?

PROBLEMAS DESAFIADORES

19. Você tem as seguintes leituras de seqüências de um clone genômico do genoma de *Homo sapiens*:

Leitura 1: ATGCGATCTGTGAGCCGAGTCTTTA

Leitura 2: AACAAAAATGTTGTTATTTTATTTTCAGATG

Leitura 3: TTCAGATGCGATCTGTGAGCCGAG

Leitura 4: TGTCTGCCATTCTTAAAAACAAAAATGT

Leitura 5: TGTTATTTTATTTTCAGATGCGA

Leitura 6: AACAAAAATGTTGTTATT

- Use as seis seqüências de leitura para criar uma seqüência contig dessa parte do genoma de *H. sapiens*.
 - Traduza a seqüência contig em todas as leituras possíveis.
 - Vá para a página BLAST do National Center for Biotechnology Information, ou NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>), e consulte o Apêndice B), e veja se você pode identificar o gene do qual essa seqüência é parte usando as matrizes de leitura como indagação para comparação proteína-proteína (BLASTp).
20. Algumas regiões mensuráveis de cromossomos diferentes do genoma humano são mais de 99% idênticas em nucleotídeos umas em relação às outras. Essas regiões foram ignoradas na produção do rascunho da seqüência do genoma humano devido a seu alto grau de similaridade. Das técnicas discutidas neste capítulo, quais permitem que os pesquisadores do genoma identifiquem a existência de tais regiões duplicadas?
21. Alguns éxons do genoma humano são bem pequenos (menos de 75 pb de tamanho). A identificação de tais "microéxons" é difícil porque as distâncias são muito curtas para um uso confiável de identificação de ORF ou tendenciosidade de códon para determinar se pequenas seqüências genômicas são verdadeiramente parte de um mRNA e um polipeptídeo. Que técnicas para "encontrar genes" podem ser usadas para tentar avaliar se determinada região de 75 pb constitui um éxon?
22. Você está estudando proteínas tendo papéis na tradução no camundongo. Pela análise BLAST das proteínas previstas do genoma de camundongo, você identifica um conjunto de genes de camundongo que codifica proteínas com seqüências similares àquelas dos conhecidos fatores eucarióticos de início de tradução. Você está interessado em determinar os fenótipos associados a mutações de perda de função desses genes.
- Você usaria genética direta ou reversa para identificar essas mutações?
 - Esquematize resumidamente dois enfoques diferentes que você pode usar para procurar fenótipos de perda de função em um desses genes.
23. Todo o genoma da levedura *Saccharomyces cerevisiae* foi seqüenciado. Esse seqüenciamento levou à identificação de todas as matrizes de leitura abertas (ORF, seqüências

do tamanho do gene com sinais apropriados de início e término da tradução) no genoma. Algumas dessas ORF são genes previamente conhecidos com funções estabelecidas; entretanto, o restante são matrizes de leitura incógnitas (URF). Para deduzir as possíveis funções das URF, elas estão sendo sistematicamente, uma de cada vez, convertidas em alelos nulos por técnicas de nocaute *in vitro*. Os resultados são os seguintes:

15% são letais quando nocauteadas.

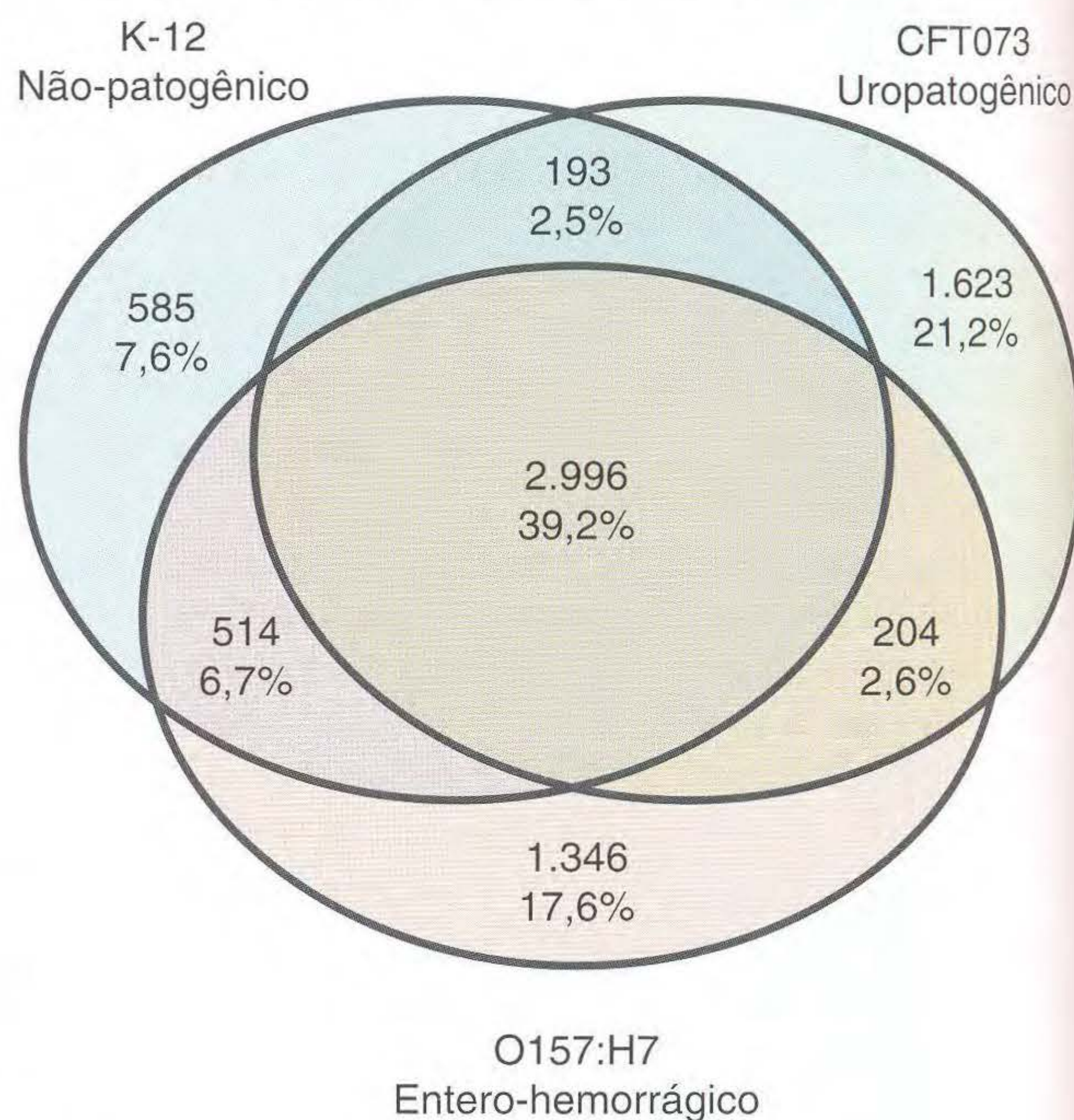
25% apresentam algum fenótipo mutante (alterado morfológicamente, com nutrição alterada, e assim por diante).

60% não apresentam fenótipo mutante detectável e assemelham-se ao tipo selvagem.

Explique a possível base genético-molecular dessas três categorias mutantes, inventando exemplos quando possível.

24. Linhagens diferentes de *E. coli* são responsáveis por infecções entero-hemorrágicas e do trato urinário. Com base nas diferenças entre a linhagem benigna K-12 e a linhagem entero-hemorrágica O157:H7, você poderia prever que existem óbvias diferenças genômicas:

- Entre K-12 e linhagens uropatogênicas?
- Entre O157:H7 e linhagens uropatogênicas?



Total de proteínas = 7.638

2.996 (39,2%) em todas as três linhagens

911 (11,9%) em duas das três linhagens

3.554 (46,5%) em uma das três linhagens

- O que pode explicar as diferenças observadas par a par no conteúdo do genoma?
- Como pode ser testada a função de genes específicos de linhagens?

EXPLORANDO GENOMAS Um Tutorial de Bioinformática Baseado na Web

Introdução a Bancos de Dados Genômicos

Onde um pesquisador pode ter informações sobre um gene? Os bancos de dados genéticos integrados são mantidos por várias organizações privadas e governamentais. No primeiro tutorial de genômica no Web site www.whfreeman.com/iga9e, você será introduzido aos recursos disponíveis no National Center for Biotechnology Information (NCBI) em Washington, D.C.

Aprendendo a Usar o ENTREZ

O programa ENTREZ no NCBI é uma ferramenta de pesquisa integrada que liga uma variedade de tipos diferentes de conteúdos. No tutorial Genomics, no Web site www.whfreeman.com/iga9e, você pode usá-lo para procurar o gene da distrofina associado à distrofia muscular e encontrar referências a pesquisas na literatura, a sequência do gene, bem como os domínios conservados, o gene equivalente para uma variedade de organismos que não os humanos e o mapa cromossômico de sua localização.

Aprendendo a Usar o BLAST

Para comparar uma sequência de proteína com outra, frequentemente usamos um programa de computação chamado BLAST. Esse programa nos permite usar uma sequência de proteínas para pesquisar e encontrar sequências de outros organismos que são similares a ela. No tutorial Genomics no Web site www.whfreeman.com/iga9e, você pode correr o BLAST em uma pequena proteína, insulina (veja o Capítulo 11), e uma grande e complexa, distrofina.

Uso de BLAST para Comparar Sequências de Ácidos Nucleicos

O algoritmo BLAST também é capaz de procurar sequências de ácidos nucleicos e compará-las. No tutorial Genomics no

Web site www.whfreeman.com/iga9e, você vai descobrir que essa comparação de sequências de RNA transportador entre espécies é um bom modo para explorar essa capacidade.

Aprendendo a Usar o PubMed

O PubMed fornece um banco de dados da literatura científica mundial. No tutorial Genomics no Web site www.whfreeman.com/iga9e, você vai aprender a pesquisar a literatura para encontrar a primeira sequência de gene repórter e publicações subsequentes demonstrando a função do gene.

Aglomerados de Grupos Ortólogos

À medida que crescem as sequências de bancos de dados, cada vez mais achamos genes similares em espécies diferentes. Esses ortólogos podem ser analisados para investigar seu grau de conservação e sua distribuição na árvore filogenética. No tutorial Genomics, no Web site www.whfreeman.com/iga9e, você aprenderá como fazer tais investigações usando banco de dados COG, os quais contêm informações sobre a conservação e distribuição de ortólogos obtidos de genomas totalmente sequenciados.

Análise de Genoma Total

Hoje temos muitos genomas totalmente sequenciados com que lidar. Sua disponibilidade permite a análise por computador como a usada para analisar os bancos de dados COG. Ela também nos permite criar experimentos para testar quais genes funcionam em quais processos e como os vários produtos gênicos interagem uns com os outros. No tutorial Genomics no Web site www.whfreeman.com/iga9e, você verá como investigar essas questões no âmbito de todo o genoma usando técnicas tais como deleção gênica para examinar o fenótipo de perda de função ou métodos para investigar a interação proteína-proteína em grande escala.

O Genoma Dinâmico: Elementos de Transposição

14



Grãos em uma espiga de milho. Os grãos manchados nessa espiga de milho resultam da interação de um elemento genético móvel (um elemento de transposição) com um gene de milho cujo produto é necessário para a pigmentação. [Cliff Weil and Susan Wessler.]

Nasceu um menino com uma doença que torna seu sistema imune não-funcional. Os testes diagnósticos determinaram que ele tem um distúrbio genético recessivo chamado SCID (doença de imunodeficiência combinada grave), mais comumente conhecida como *doença do menino da bolha*. Essa doença é causada por uma mutação no gene que codifica a enzima sanguínea adenosina desaminase (ADA). Como resultado da perda dessa enzima, as células precursoras que dão origem a um dos tipos celulares do sistema imune estão faltando. Como esse menino não tem a habilidade de combater uma infecção, ele tem que viver em um ambiente completamente isolado e estéril, isto é, uma bolha na qual o ar é filtrado para ser estéril (Figura 14.1). Nenhuma terapia farmacêutica ou outra terapia convencional está disponível para tratar essa doença. Dar ao menino um transplante contendo as células precursoras de outra pessoa não funcionaria, na grande maioria dos casos, porque uma correspondência tissular precisa entre o doador e o paciente é extremamente rara. Conseqüentemente, as células doadoras terminariam criando uma resposta imune contra os próprios tecidos do menino (doença enxerto-versus-hospedeiro).

Nas últimas duas décadas, foram desenvolvidas técnicas que oferecem a possibilidade de um tipo diferente de terapia de transplante, a **terapia gênica**, que pode ajudar as pessoas com SCID e outras doenças incuráveis. Com relação à SCID, um gene ADA normal é “transplantado” para células do sistema imunológico do paciente, permitindo assim sua sobrevivência e função normal. Nas primeiras triagens de terapia gênica humana, os cientistas modificaram

Perguntas Fundamentais

- Por que os elementos de transposição foram primeiro descobertos geneticamente em milho, mas primeiro isolados molecularmente de *E. coli*?
- Como os elementos de transposição participam da dispersão de bactérias resistentes a antibióticos?
- Por que os elementos de transposição são classificados como transposons de DNA ou transposons de RNA?
- Como os humanos podem sobreviver, considerando que até 50% do genoma humano é derivado de elementos de transposição?

Tópicos

- 14.1 Descoberta dos elementos de transposição em milho
- 14.2 Elementos de transposição em procariontes
- 14.3 Elementos de transposição em eucariontes
- 14.4 O genoma dinâmico: mais elementos de transposição do que se imaginava

Um menino com SCID



FIG. 14.1 Um paciente com SCID deve viver em uma bolha de proteção. [UPI/Bettmann/CORBIS.]

um tipo de vírus chamado de retrovírus no laboratório (“construído”) de modo que ele pudesse se inserir a um gene ADA normal nos cromossomos das células imunes colhidas de pacientes com SCID. Neste capítulo, você verá que os retrovírus têm muitas propriedades biológicas em comum com um tipo de elemento móvel chamado de *retrotransposon*, que popula nosso genoma e os genomas da maioria dos eucariontes. As lições aprendidas sobre o comportamento dos retrotransposons e outros elementos móveis de organismos-modelo, tais como leveduras, são fontes de valiosos *insights* sobre o *design* de uma nova geração de rearranjos biológicos para a terapia de genes humanos.

Começando nos anos 1930, os estudos genéticos de milho deram resultados que superaram muito o quadro da genética clássica dos genes residindo apenas em loci fixos dos cromossomos. A literatura de pesquisa começou a ter relatos sugerindo que alguns elementos genéticos presentes nos cromossomos podiam, de algum modo, mover-se de um local para outro. Esses achados foram vistos com ceticismo por muitos anos, mas hoje está claro que tais elementos móveis são muito dispersos na natureza.

Uma variedade de nomes coloridos (alguns que ajudam a descrever suas respectivas propriedades) foram aplicados a esses elementos genéticos: elementos controladores, genes saltadores, genes móveis, elementos móveis e transposons. Aqui usaremos os termos *elementos transponíveis* e *elementos móveis*, que envolvem toda a família de tipos. Elementos de transposição podem mover-se para novas posições dentro do mesmo cromossomo ou mesmo para um cromossomo diferente. Eles foram detectados geneticamente em organismos-modelo, tais como *E. coli*, milho, leveduras e *Drosophila*, por meio de mutações que eles produzem quando se inserem e inativam genes.

O sequenciamento do DNA de genomas de uma variedade de micróbios, plantas e animais indica que os elementos de transposição existem em praticamente todos os organismos.

Surpreendentemente, eles são de longe o maior componente do genoma humano, correspondendo a quase 50% de nossos cromossomos. A despeito de sua abundância, o papel genético normal desses elementos não é conhecido com certeza.

Em seus estudos, os cientistas são capazes de explorar a habilidade dos elementos de transposição em inserir-se em novos locais no genoma. Os elementos de transposição construídos no tubo de ensaio são instrumentos valiosos, tanto em procariontes quanto em eucariontes, para mapeamento genético, criação de mutantes, clonagem de genes e, mesmo, para a produção de organismos transgênicos (veja o Capítulo 20). Vamos reconstruir algumas das etapas na evolução de nossa presente compreensão dos elementos de transposição. Ao fazermos isso, descobriremos os princípios que orientam essas fascinantes unidades genéticas.

► 14.1 Descoberta dos Elementos de Transposição em Milho

Os experimentos de McClintock: o elemento Ds

Nos anos de 1940, Barbara McClintock fez uma incrível descoberta enquanto estudava os grãos coloridos do chamado milho indiano, conhecido como milho (veja o box Organismo-modelo adiante). O milho tem 10 cromossomos, numerados do maior (1) até o menor (10). Ao analisar a quebra dos cromossomos de milho, McClintock notou um fenômeno incomum. Ela descobriu que, em uma linhagem de milho, o cromossomo 9 quebrava-se muito frequentemente em determinado local ou locus (Figura 14.2). Ela determinou que a quebra do cromossomo nesse locus era devida à presença de dois fatores genéticos. Um fator que ela chamou de **Ds** (para **Dissociação**) estava situado no local da quebra. Outro, um fator genético não-ligado, era necessário para “ativar” a quebra do cromossomo 9 no locus Ds. Assim, McClintock chamou esse segundo fator de **Ac** (de **Ativador**).

McClintock começou a suspeitar que **Ac** e **Ds** eram de fato elementos genéticos móveis quando ela achou impossível mapear **Ac**. Em algumas plantas, ele era mapeado em uma posição; em outras plantas da mesma linhagem, era mapeado em posições diferentes. Como essa variável de mapeamento

O elemento de transposição Ds ajuda a causar quebras

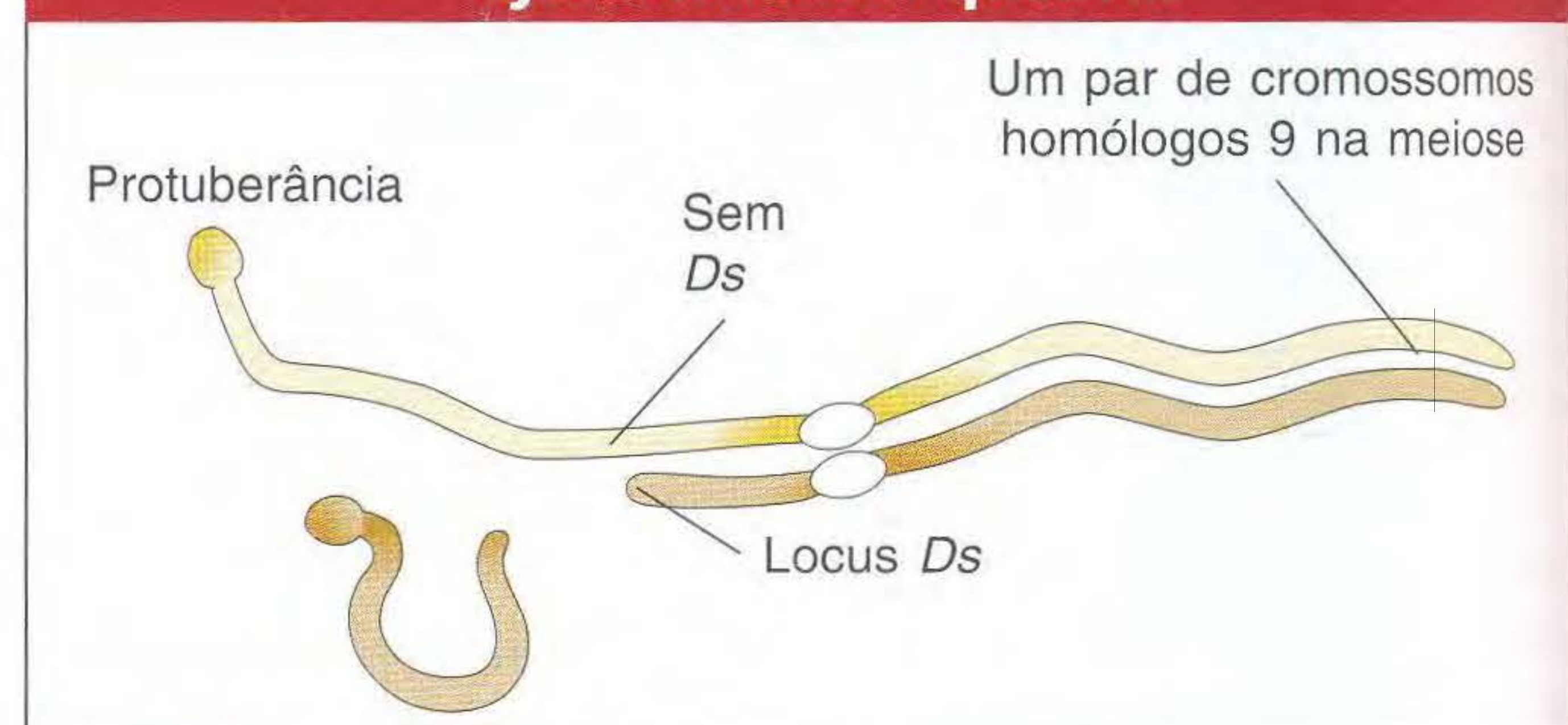


FIG. 14.2 O cromossomo 9 de milho quebra no locus Ds, onde o elemento de transposição Ds foi inserido.



Organismo-modelo Milho

O milho, *Zea mays*, é um membro da família das gramíneas. As gramíneas também incluem o arroz, o trigo e a cevada, e são a fonte mais importante de calorias para a humanidade. O milho foi domesticado da gramínea selvagem teosinto pelos americanos nativos no México e América Central, e foi primeiro introduzido na Europa por Colombo em seu retorno do Novo Mundo.

Nos anos de 1920, Rollins A. Emerson criou um laboratório na Cornell University para estudar a genética das características do milho, incluindo a cor do grão, ideais para a análise genética. Além disso, a separação física de flores masculinas e femininas na borla e espiga, respectivamente, permitiu fazer cruzamentos genéticos controlados fáceis de realizar. Entre os geneticistas de destaque atraídos pelo laboratório de Emerson, estavam Marcus Rhoades, Barbara McClintock e George Beadle (veja o Capítulo 6). Antes do advento da Biologia Molecular e da ascensão dos microrganismos como organismos-modelo, os geneticistas fizeram análises microscópicas de cromossomos e correlacionaram seu comportamento com a segregação das características. Os grandes cromossomos de paquíteno do milho e a glândula salivar de *Drosophila* os tornaram os organismos de escolha para a análise citogenética. Os resultados desses estudos iniciais levaram a uma compreensão do comportamento cromossômico durante a meiose e a mitose, incluindo eventos como a recombinação e as consequências de quebras cromossômicas, tais como inversões, translocações e duplicação.

O milho ainda serve como organismo genético modelo. Os biólogos moleculares continuam a explorar seus belos

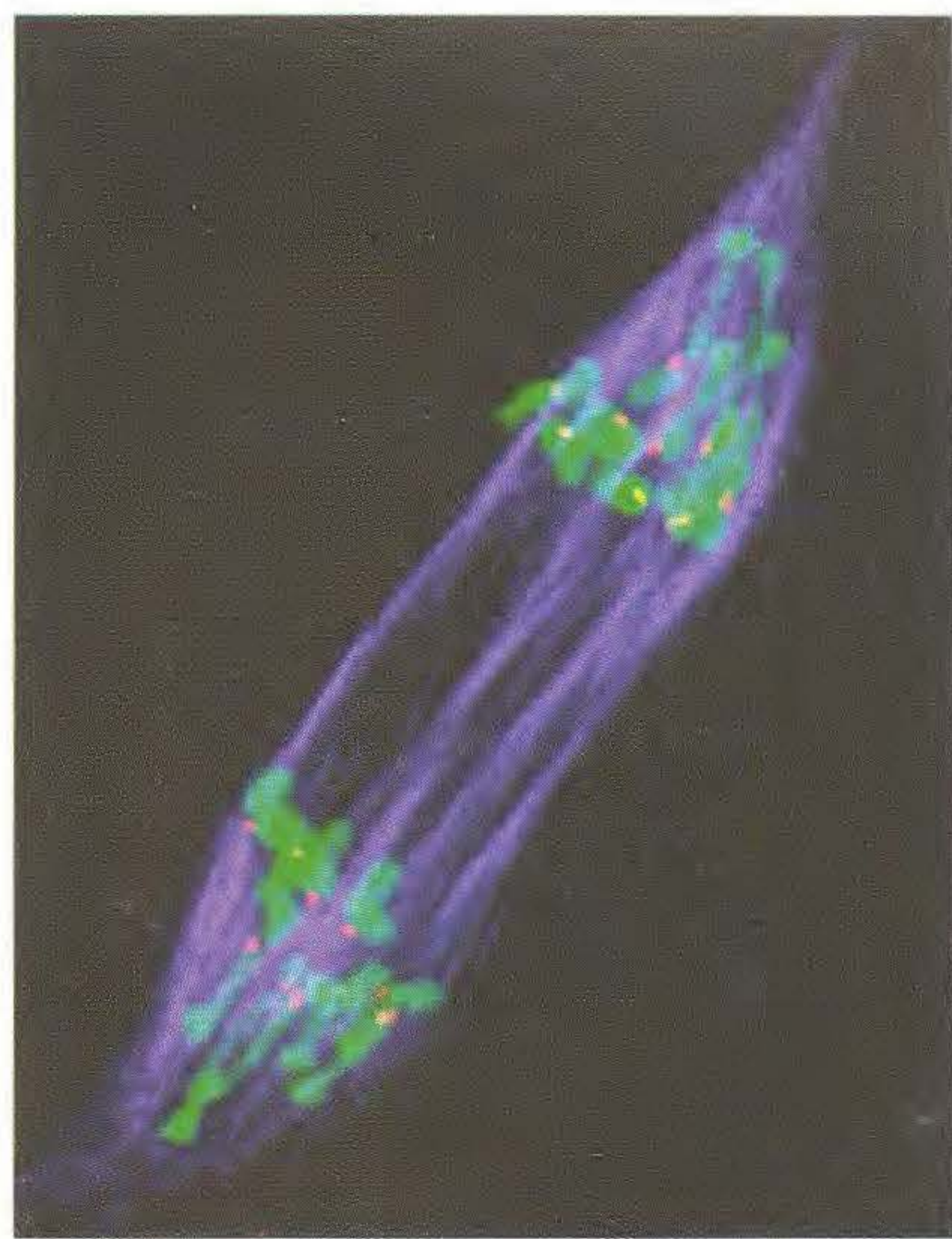


O laboratório de milho de Rollins A. Emerson na Cornell University, 1929. Da esquerda para a direita: Charles Burnham, Marcus Rhoades, Rollins Emerson e Barbara McClintock. Agachado está George Beadle. Tanto McClintock quanto Beadle (veja o Capítulo 6) receberam o Prêmio Nobel. [Cortesia do Department of Plant Breeding, Cornell University.]

cromossomos do paquíteno com novas sondas de anticorpos (veja a Foto *b* em seguida) e usaram seus elementos transponíveis geneticamente bem caracterizados como ferramentas para identificar e isolar genes importantes.



(a)



(b)

Análise dos cromossomos de milho, antes e agora. Os cromossomos de milho são grandes e facilmente visíveis ao microscópio óptico. (a) Uma imagem tomada por Marcus Rhoades (1952). (b) Essa imagem é comparável à da parte *a*, porém o fuso é mostrado em azul (corado com anticorpos para tubulina), os centrômeros são mostrados em vermelho (corados com anticorpos a uma proteína associada ao centrômero), e os cromossomos são mostrados em verde. [(a) De M.M. Rhoades, "Preferential Segregation in Maize," in J.W. Gowen, Ed. *Heterosis*, pp. 66-80. Iowa State College Press, 1952. (b) De R.K. Dawe, L. Reed, H.-G. Yu, M.G. Muszynski, and E.N. Hiatt. "A Maize Homolog of Mammalian CENPC Is a Constitutive Component of the Inner Kinetochores," *Plant Cell* 11, 1999, 1227-1238.]

não era mais que uma curiosidade, raros grãos com fenótipos marcadamente diferentes podem ser derivados da linhagem original que tinham quebras freqüentes no cromossomo 9. Um desses fenótipos era um raro grão incolor contendo pontos de pigmentos.

A Figura 14.3 compara o fenótipo da linhagem com quebras cromossômicas com o fenótipo de uma dessas linhagens derivadas. Para a linhagem com quebras cromossômicas, um cromossomo que se quebra em ou perto de *Ds* perde sua ponta contendo os alelos tipo selvagem dos genes *C*, *Sh* e *Wx*. No exemplo mostrado na Figura 14.3a, vê-se uma quebra ocorrida em uma única célula, que se dividiu mitoticamente para produzir o grande setor de tecido mutante (*c sh wx*). A quebra pode ocorrer muitas

vezes em um único grão, mas cada setor de tecido apresentará a perda de expressão de todos os três genes. Em contraste, cada novo derivado afeta a expressão apenas de um único gene. Um derivado que afeta a expressão apenas do gene de pigmento *C* é mostrado na Figura 14.3b. Nesse exemplo, os pontos pigmentados aparecem sobre um fundo incolor da espiga. Embora a expressão de *C* esteja alterada desse modo estranho, a expressão de *Sh* e *Wx* era normal e o cromossomo 9 não apresentava mais quebras freqüentes.

Para explicar os novos derivados, McClintock formulou a hipótese de que *Ds* tinha se movido de um sítio perto do centrômero para o gene *C* situado perto da

Fenótipos incomuns são causados pelo elemento de transposição *Ds*

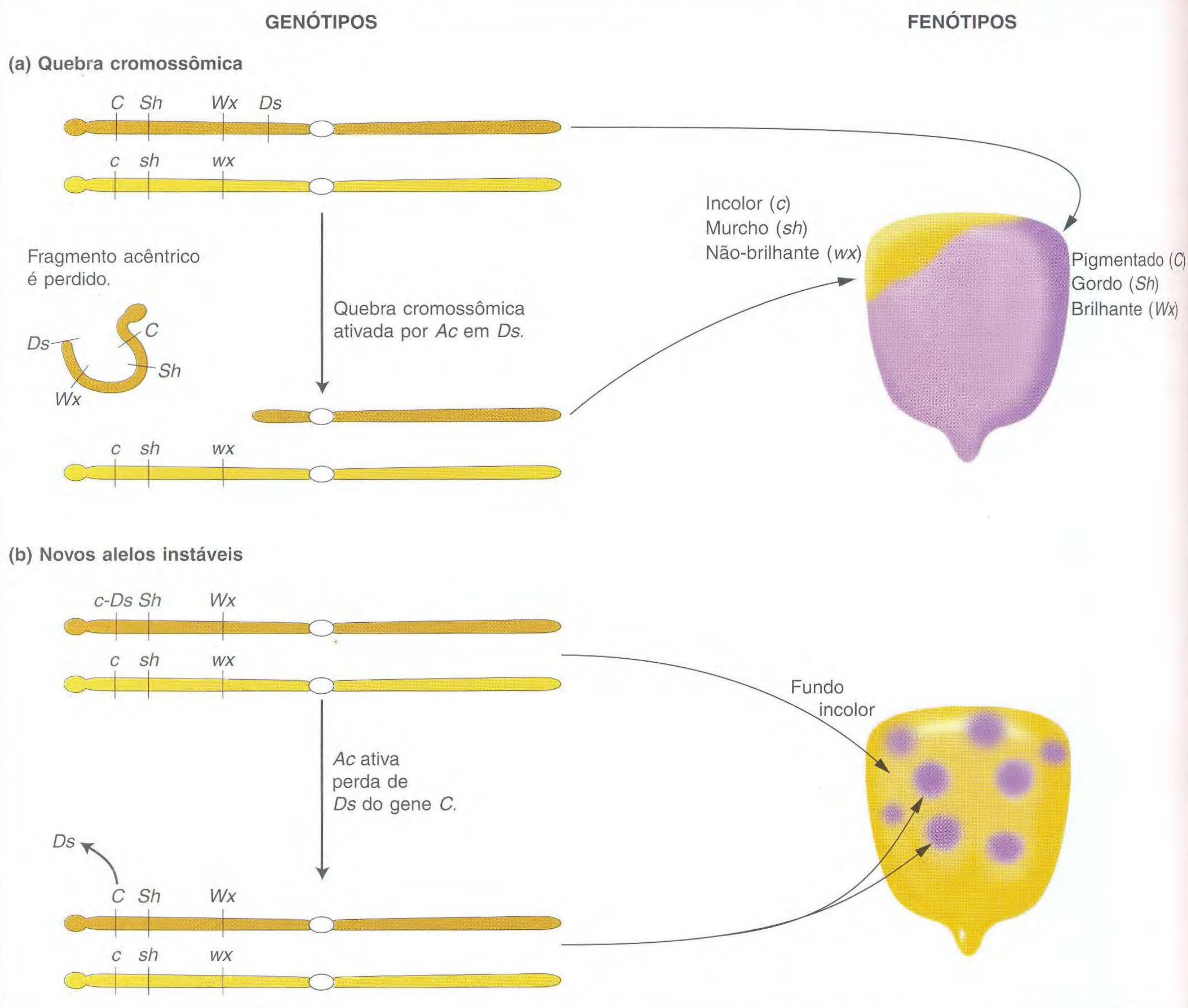


FIG. 14.3 Novos fenótipos em milho são produzidos pelo movimento do elemento de transposição *Ds* no cromossomo 9. (a) Um fragmento de cromossomo é perdido pela quebra no locus *Ds*. Os alelos recessivos do cromossomo homólogo são expressos, produzindo o setor incolor no grão. (b) A inserção de *Ds* no gene *C* (em cima) cria células de grãos incolores. A excisão de *Ds* do gene *C*, pela ação de *Ac* nas células e seus descendentes mitóticos, permite que a cor seja novamente expressa, produzindo o fenótipo pontilhado.

ponta telomérica. Em sua nova localização, *Ds* impede a expressão de *C*. A inativação do gene *C* explica as partes incolores dos grãos, mas o que explica o aparecimento dos pontos pigmentados? O grão pigmentado é um exemplo de um **fenótipo instável**. McClintock concluiu que tais fenótipos instáveis resultavam do movimento ou transposição de *Ds* para longe do gene *C*. Isto é, os grãos começam o desenvolvimento com um gene *C* que foi mutado pela inserção de *Ds*. Entretanto, em algumas células do grão, *Ds* deixa o gene *C*, permitindo que o fenótipo mutante reverta ao tipo selvagem e produza um pigmento na célula original e em todas as suas descendentes mitóticas. Existem grandes pontos de cor quando *Ds* sai do gene *C* mais cedo no desenvolvimento do grão (pois há mais descendentes mitóticos), enquanto existem pequenos pontos quando *Ds* deixa o gene *C* mais tarde no desenvolvimento do grão. Fenótipos mutantes instáveis que reverterem ao tipo selvagem são um indício da participação dos elementos móveis.

Elementos autônomos e não-autônomos

Qual a relação entre *Ac* e *Ds*? Como eles interagem com genes e cromossomos para produzir esses fenótipos interessantes e incomuns? Essas questões foram respondidas por análise genética posterior. As interações entre *Ds*, *Ac* e o pigmento do gene *C* são usadas como um exemplo na Figura 14.4. Nessa figura, *Ds* é mostrado como um pedaço do DNA que tem inativado o gene *C* pela inserção em sua região codificante. O alelo levando o inserto é chamado de *c-mutável(Ds)* ou, abreviadamente, *c-m(Ds)*. Uma linhagem com *c-m(Ds)* e nenhum *Ac* tem grãos incolores porque *Ds* não pode mover-se; ele está parado no gene *C*. Uma linhagem com *c-m(Ds)* e *Ac* tem grãos pontilhados porque *Ac* ativa *Ds* em algumas células e deixa o gene *C*, restaurando assim a função do gene. O elemento que sai é dito **excisado** do cromossomo ou **transposto**.

Foram isoladas outras linhagens nas quais o elemento *Ac* inseriu-se no gene *C* [chamado de *c-m(Ac)*]. Ao contrário do alelo *c-m(Ds)*, que é instável apenas quando *Ac* está no genoma, *c-m(Ac)* é sempre instável. Além disso, McClintock descobriu que, em raras ocasiões, um alelo do tipo *Ac* pode ser transformado em um alelo do tipo *Ds*. Essa transformação era devida à geração espontânea do elemento *Ds* a partir do elemento *Ac* inserido. Em outras palavras, *Ds* é, muito provavelmente, uma versão mutada incompleta do próprio *Ac*.

Vários sistemas como *Ac/Ds* foram descobertos por McClintock e outros geneticistas que trabalhavam com milho. Dois outros sistemas são *Dotted (Dt)*, descoberto por Marcus Rhoades, e *Suppressor/mutator (Spm)*, descoberto independentemente por McClintock e Peter Peterson, que o chamou de *Enhancer/Inhibitor (En/In)*. Além disso, você verá, nas seções seguintes, que elementos com comportamento genético similar foram isolados em bactérias, plantas e animais.

O comportamento genético comum desses elementos levou os geneticistas a propor novas categorias para todos os elementos. *Ac* e elementos com propriedades genéticas

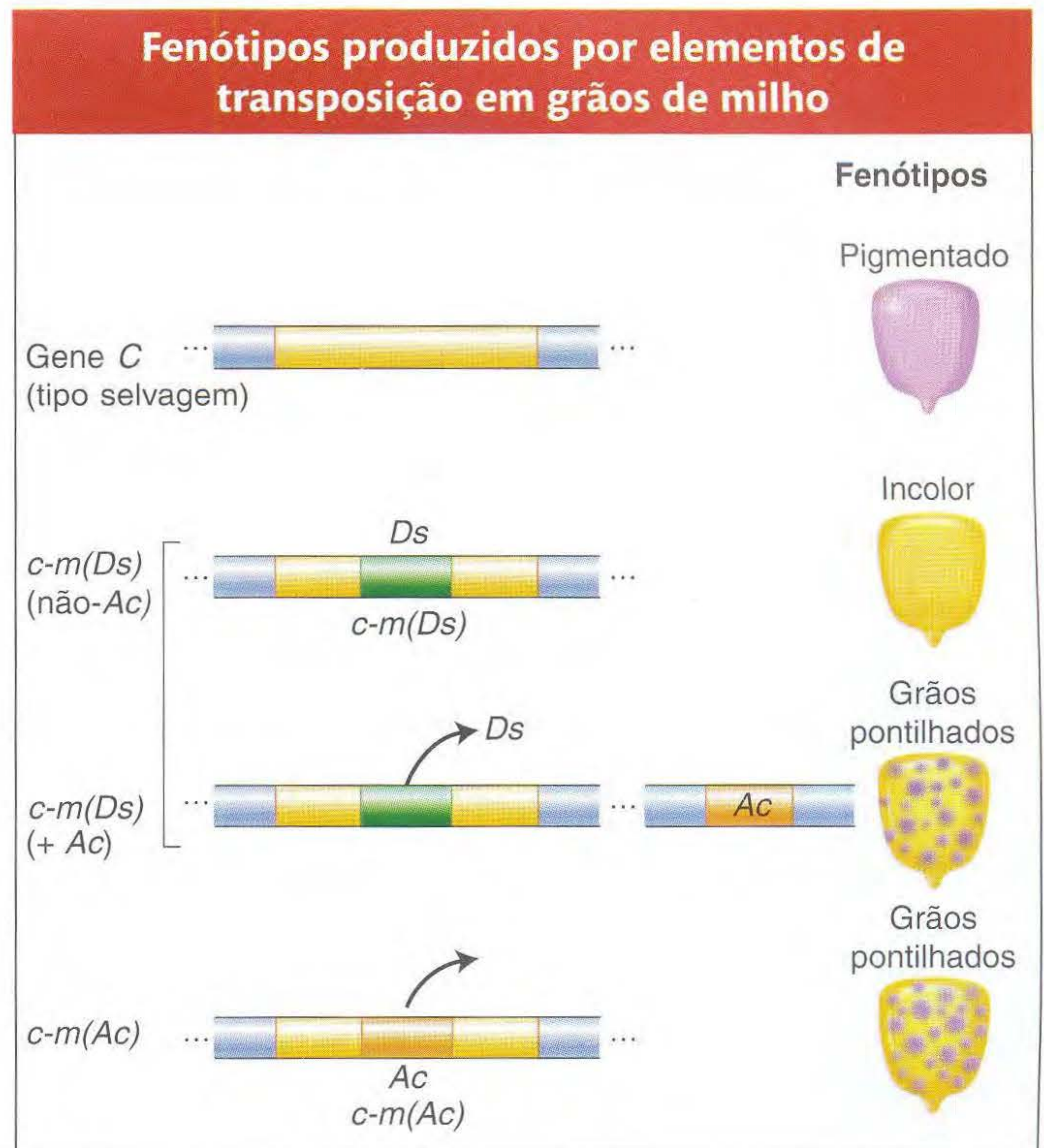


FIG. 14.4 O pontilhado do grão é controlado pela inserção e excisão dos elementos *Ds* ou *Ac* no gene *C* que controla o pigmento.

similares hoje são chamados de **elementos autônomos** porque não necessitam de outros elementos para sua mobilidade. Similarmente, *Ds* e elementos com propriedades genéticas similares são chamados de **elementos não-autônomos**. Um elemento *família* é composto de um ou mais elementos autônomos e membros não-autônomos que podem ser mobilizados. Os elementos autônomos codificam a informação necessária para seu próprio movimento e para o movimento de elementos não-autônomos não-ligados no genoma. Como os elementos não-autônomos não codificam as funções necessárias para seu próprio movimento, eles não podem mover-se a menos que um elemento autônomo de sua família esteja presente em outra parte do genoma.

A Figura 14.5 mostra um exemplo dos efeitos dos transposons em boca-de-leão.

Os elementos de transposição em milho podem inativar um gene no qual residem, causar quebras cromossômicas e transpor-se para novos locais dentro do genoma. Os elementos autônomos podem fazer essas funções sem ajuda; os elementos não-autônomos só podem transpor-se com a ajuda de um elemento autônomo em outra parte do genoma.

Elementos transponíveis: só em milho?

Embora os geneticistas tenham aceitado a descoberta de McClintock dos elementos transponíveis em milho, muitos relutaram em considerar a possibilidade de que elementos

Elementos de transposição funcionando em boca-de-leão



FIG. 14.5 O mosaicismo é causado pela excisão de elementos de transposição em boca-de-leão (*Antirrhinum*). A inserção do elemento de transposição perturba a produção de pigmento, resultando em flores brancas. A excisão do elemento de transposição restaura a produção do pigmento, resultando em setores de tecido floral vermelho. [Foto de Rosemary Carpenter e Enrico Coen.]

similares residam nos genomas de outros organismos. Sua existência em todos os organismos significaria que os genomas são herdavelmente instáveis e dinâmicos. Essa visão era inconsistente com o fato de que os mapas genéticos de membros da mesma espécie fossem iguais. Se os genes podem ser mapeados geneticamente em um local cromossômico preciso, esse mapeamento não indica que eles não estão se movendo?

Como McClintock era uma geneticista muito respeitada, seus resultados foram explicados dizendo que milho não era um organismo natural; é uma planta que é produto de seleção humana e domesticação. Essa visão foi mantida até os anos de 1960, quando os primeiros elementos de transposição foram isolados do genoma de *E. coli* e estudados ao nível da sequência de DNA. Os elementos de transposição foram subsequenteiramente isolados de genomas de muitos organismos, incluindo *Drosophila* e leveduras. Quando se tornou aparente que os elementos de transposição são um componente significativo dos genomas da maioria e, talvez, de todos os organismos, Barbara McClintock foi reconhecida por sua descoberta importante e recebeu o Prêmio Nobel de 1983 em Medicina ou Fisiologia.

► 14.2 Elementos de Transposição em Procariontes

A descoberta genética dos elementos de transposição levou a muitas dúvidas sobre com o que tais elementos poderiam parecer no que concerne a sequência de DNA e como eles eram capazes de mover-se de um local para outro no geno-

ma. Todos os organismos os têm? Todos os elementos são similares ou existem classes diferentes de elementos de transposição? Se existem muitas classes de elementos, eles podem coexistir em um genoma? O número de elementos de transposição no genoma varia de espécie para espécie? A natureza molecular dos elementos genéticos de transposição foi primeiro compreendida em bactérias. Portanto, consideraremos essa história examinando os estudos originais feitos com procariontes.

Seqüências bacterianas de inserção

As seqüências de inserção, ou **elementos de seqüência de inserção (IS)**, são segmentos de DNA bacteriano que podem mover-se de uma posição em um cromossomo para uma posição diferente no mesmo cromossomo ou em um cromossomo diferente. Quando os elementos IS aparecem no meio de genes, eles interrompem a seqüência codificante e inativam a expressão desse gene. Devido a seu tamanho e alguns casos de presença de sinais de término de transcrição e tradução no elemento IS, os elementos IS também podem bloquear a expressão de outros genes no mesmo óperon se esses genes são posteriores ao sítio de inserção. Os elementos IS foram primeiro encontrados em *E. coli* no óperon *gal*, um grupo de três genes que participam do metabolismo do açúcar galactose.

Identificação de pequenos elementos IS Vários mutantes *gal⁻* de *E. coli* foram encontrados contendo grandes inserções de DNA no óperon *gal*. Esse achado levou naturalmente à pergunta seguinte: Os segmentos de DNA que se inserem nos genes são meros fragmentos de DNA ou são entidades genéticas distintas? A resposta para essa pergunta veio dos resultados dos experimentos de hibridização mostrando que muitas mutações de inserção diferentes são causadas por um pequeno grupo de seqüências de inserção. Esses experimentos são feitos com o uso do fago λ *dg**gal* que contém o óperon *gal⁻* de várias linhagens mutantes *gal* isoladas. As partículas individuais de fago das linhagens são isoladas, e seu DNA é usado para sintetizar RNA radioativo *in vitro*. Alguns fragmentos desse RNA são encontrados hibridizando-se ao DNA de outras mutações *gal⁻* contendo grandes inserções de DNA, mas não com DNA tipo selvagem. Esses resultados foram interpretados como significando que mutantes *gal* isolados independentemente contêm o mesmo pedaço extra de DNA. Esses fragmentos particulares de RNA também se hibridizam com o DNA de outros mutantes contendo as inserções IS em outros genes, mostrando que o mesmo trecho de DNA pode inserir-se em locais diferentes no cromossomo bacteriano.

Com base em seus padrões de hibridização cruzada, vários elementos IS distintos foram identificados. Uma seqüência, chamada IS1, é o segmento de 800 pb identificado em *gal*. Outra seqüência, chamada IS2, tem 1.350 pb de tamanho. Embora os elementos IS difiram em seqüência de DNA, eles têm várias características em comum. Por exemplo, todos os elementos IS codificam uma proteína, chamada **transposase**, que é uma enzima necessária para o movimento dos elementos IS de um sítio no cromossomo para outro.

Além disso, todos os elementos IS começam e terminam com seqüências curtas de repetições invertidas que são necessárias para sua mobilidade. A transposição de elementos IS e outros elementos genéticos móveis será considerada mais adiante no capítulo.

O genoma do tipo selvagem padrão de *E. coli* é rico nos elementos IS: ele contém oito cópias de IS1, cinco cópias de IS2 e cópias de outros tipos de IS não tão bem estudados. Como os elementos IS são regiões de seqüências idênticas, eles são sítios onde podem ocorrer crossings. Por exemplo, a recombinação entre o plasmídeo fator F e o cromossomo de *E. coli* para formar linhagens *Hfr* é o resultado de um único crossing entre um elemento IS situado no plasmídeo e um elemento IS situado no cromossomo (veja o Capítulo 5, Figura 5.18). Como existem vários elementos IS, o fator F pode inserir-se em vários sítios.

O genoma bacteriano contém segmentos de DNA, chamados de elementos IS, que podem mover-se de uma posição no cromossomo para uma posição diferente no mesmo cromossomo ou em um cromossomo diferente.

Transposons procarióticos

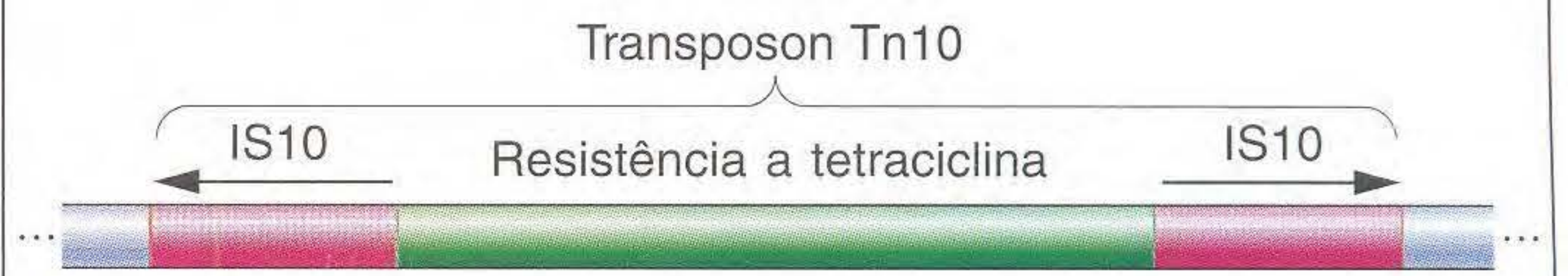
No Capítulo 5, você aprendeu sobre os **fatores R**, que são plasmídeos levando genes que codificam resistência a vários antibióticos. Esses fatores R (de resistência) são rapidamente transferidos na conjugação celular, como os fatores F em *E. coli*.

Os fatores R mostraram-se como sendo apenas os primeiros de muitos fatores similares tipo F a serem descobertos. Os fatores R foram encontrados levando muitos tipos diferentes de genes em bactérias. Em particular, os fatores R captam genes conferindo resistência a antibióticos diferentes. Como eles adquirem novas habilidades genéticas? Os genes resistentes a drogas residem em um elemento genético móvel chamado um **transposon (Tn)**. Existem dois tipos de transposons bacterianos. Os **transposons compostos** contêm uma variedade de genes que residem entre dois elementos IS quase idênticos que são orientados em direções opostas (Figura 14.6a) e, como tal, formam o que é chamado de seqüência de **repetição invertida (IR)**. A transposase codificada por um dos dois elementos IS é necessária para catalisar o movimento de todo o transposon. Um exemplo de um transposon composto é o Tn10, mostrado na Figura 14.6a. O Tn10 leva um gene que confere resistência ao antibiótico tetraciclina, e é flanqueado por dois elementos IS10 em orientação oposta. Os elementos IS que constituem transposons compostos não são capazes de se transpor devido a mutações em suas repetições invertidas.

Os **transposons simples** são flanqueados por seqüências IR, mas essas seqüências são curtas (<50 pb) e não codificam a enzima transposase que é necessária para a transposição. Assim, sua mobilidade não é devida a uma associação com elementos IS. Os transposons simples codificam sua própria transposase além de levarem genes bacterianos. Um exemplo de um transposon simples é o Tn3, mostrado na Figura 14.6b.

Características estruturais de transposons compostos e simples

(a) Transposon composto



(b) Transposon simples

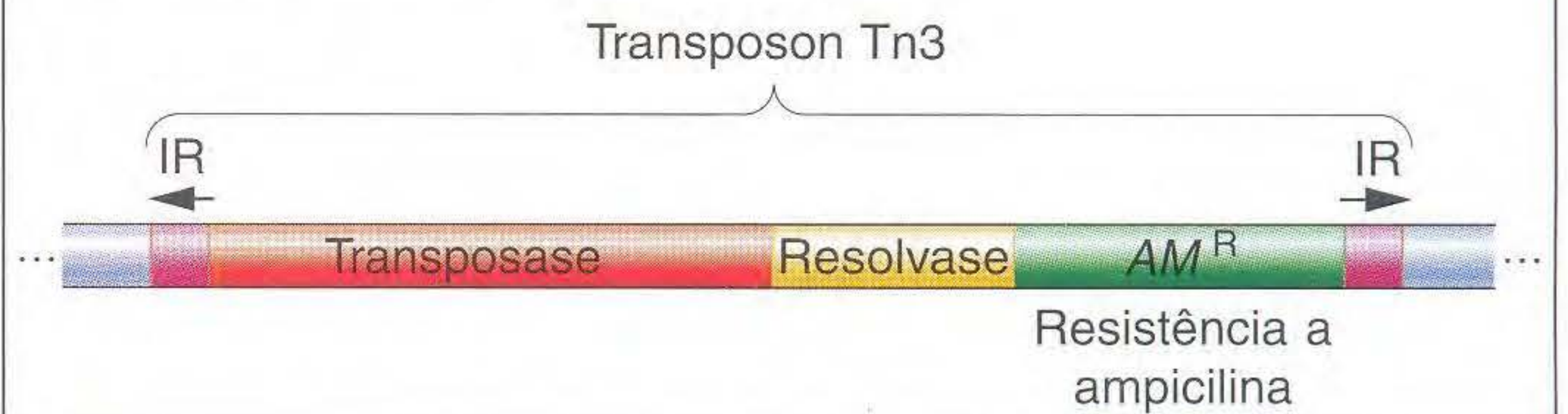


FIG. 14.6 (a) Tn10, um exemplo de um transposon composto. Os elementos IS são inseridos em orientação oposta e formam repetições invertidas (IR). Cada elemento IS leva uma transposase, mas apenas um geralmente é funcional. (b) Tn3, um exemplo de um transposon simples. Curtas repetições invertidas não contêm transposase. Os transposons simples codificam sua própria transposase. A resolvase é uma proteína que promove recombinação e resolve os cointegrados (veja Figura 14.10).

Para uma revisão, os elementos IS são seqüências móveis curtas que codificam apenas as proteínas necessárias para sua mobilidade. Os transposons compostos e os transposons simples contêm genes adicionais que conferem novas funções a células bacterianas. Sejam compostos ou simples, os transposons geralmente são chamados apenas de transposons, e transposons diferentes são chamados de Tn1, Tn2, Tn505, e assim por diante.

Um transposon pode pular de um plasmídeo para um cromossomo bacteriano ou de um plasmídeo para outro plasmídeo. Desse modo, são gerados os plasmídeos de resistência a várias drogas. A Figura 14.7 é um diagrama composto de um **plasmídeo R**, indicando os vários locais nos quais os transposons podem ser localizados. Em seguida, consideraremos a questão de como tal **transposição** ou mobilização ocorre.

Os transposons foram originalmente detectados como elementos genéticos móveis que conferem resistência a drogas. Muitos desses elementos consistem em elementos IS flanqueando um gene que codifica resistência a drogas. Essa organização promove a dispersão de bactérias resistentes a drogas facilitando o movimento do gene de resistência do cromossomo de uma bactéria resistente a um plasmídeo que pode ser conjugado a outra linhagem bacteriana (suscetível).

Mecanismo de transposição

Como já foi dito, o movimento de um elemento de transposição para um sítio em um cromossomo para outro sítio ou entre um plasmídeo e o cromossomo é mediado por uma transposase. Em uma das primeiras etapas da transposição, a transposase faz cortes desencontrados no sítio-alvo do DNA

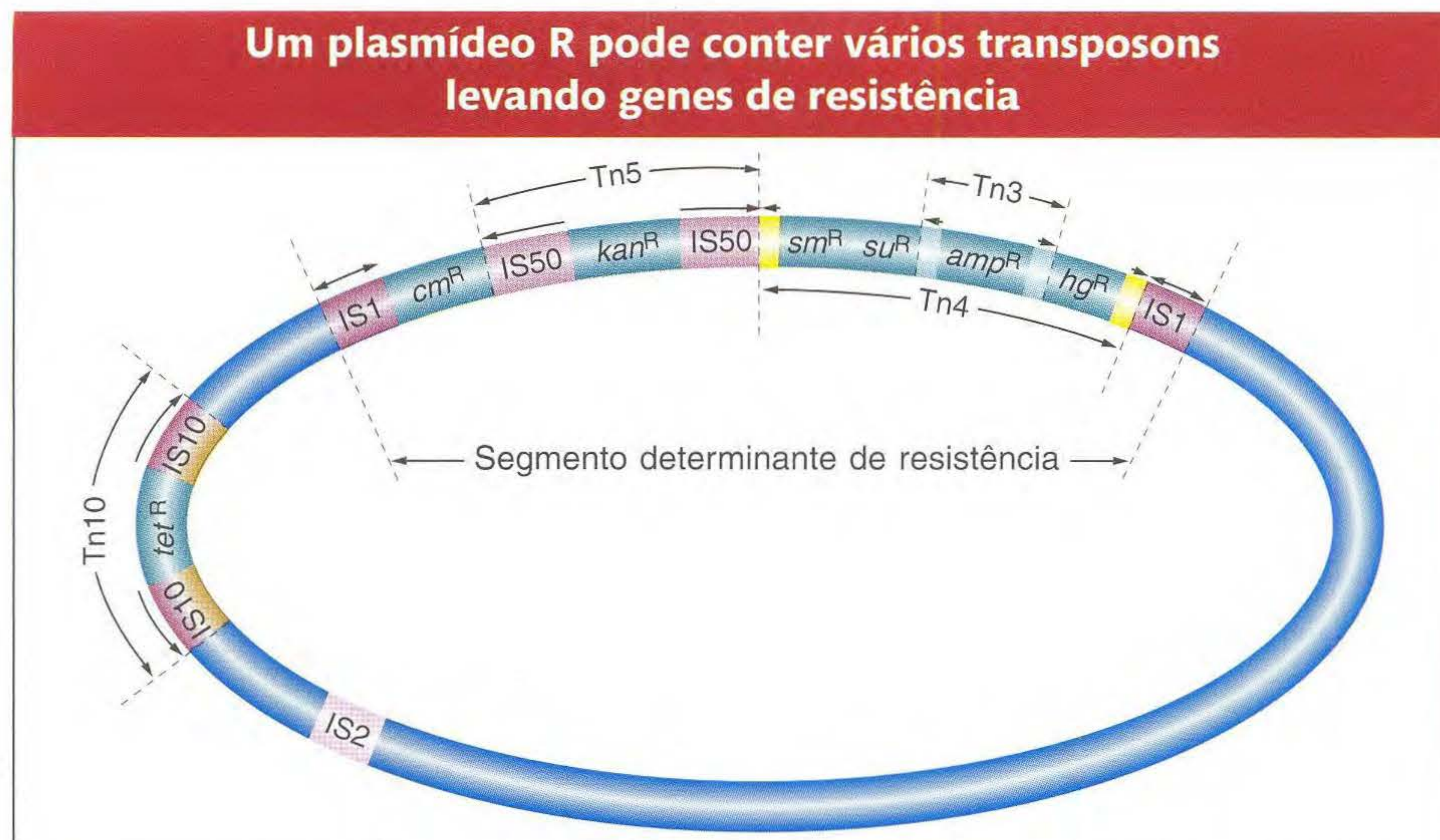


FIG. 14.7 Um mapa esquemático de um plasmídeo mostra vários transposons simples e compostos levando genes de resistência. São mostrados os genes que codificam resistência a antibióticos tetraciclina (*tet^R*), canamicina (*kan^R*), estreptomicina (*sm^R*), sulfonamida (*su^R*), e ampicilina (*amp^R*) e ao mercúrio (*hg^R*). O segmento determinante da resistência pode mover-se como um grupo de genes de resistência. Tn3 está dentro de Tn4. Cada transposon pode ser transferido independentemente. [Simplificado de S.N. Cohen and J.A. Shapiro, "Transposable Genetic Elements." Copyright 1980 by Scientific American, Inc. Todos os direitos reservados.]

(não diferentes das quebras catalisadas pelas endonucleases de restrição no arcabouço açúcar-fosfato do DNA). A Figura 14.8 mostra as etapas na integração de um transposon genérico. Nesse caso, a transposase faz um corte de cinco pares de bases. O transposon insere-se entre as pontas cortadas, e a maquinaria de reparo do DNA hospedeiro (veja o Capítulo 15) preenche o espaço unifilamentar que se projeta, usando as bases do que se projeta como um molde. Existem agora duas seqüências duplicadas, cada uma com cinco pares de bases de tamanho, nos sítios das antigas projeções. Essas seqüências são chamadas de **duplicação de sítio-alvo**. Praticamente todos os elementos de transposição (em procarionte e eucariontes) são flanqueados por uma duplicação de sítio-alvo, indicando que todos usam um mecanismo de integração similar ao mostrado na Figura 14.8. O que difere é o tamanho da duplicação; um tipo particular de elemento de transposição tem um tamanho característico para sua duplicação de sítio-alvo, tão pequeno quanto dois pares de bases para alguns elementos.

A maior parte dos elementos de transposição em procariontes (e em eucariontes) emprega um de dois mecanismos de transposição, chamados de **replicativo** e **conservativo** (não-replicativo), como ilustrado na Figura 14.9. Na via replicativa (como mostrado para Tn3), uma nova cópia do elemento de transposição é gerada no evento da transposição. Esses resultados da transposição consistem em uma cópia que aparece em um novo sítio e uma cópia que fica no sítio antigo. Na via conservativa (como mostrado para Tn10), não há replicação. Em vez disso, o elemento é removido do cromossomo ou plasmídeo e é integrado ao novo sítio. A via conservativa também é chamada de **"recortar e colar"**.

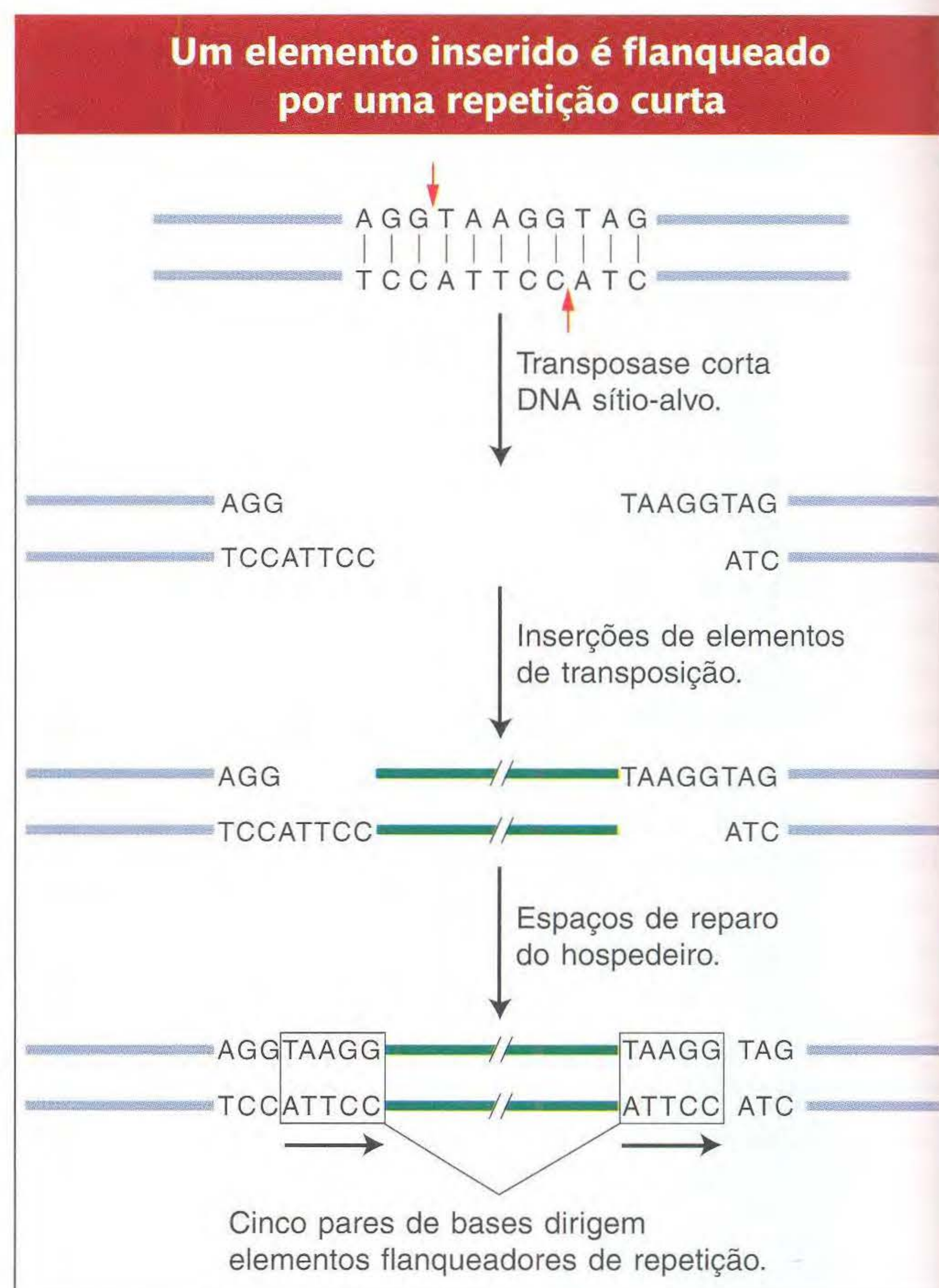


FIG. 14.8 Uma seqüência curta de DNA é duplicada no sítio de inserção do transposon. O DNA receptor é clivado em sítios desencontrados (é mostrado um corte desencontrado de 5 pb), levando à produção de duas cópias da seqüência de cinco pares de bases flanqueadoras do elemento de inserção.

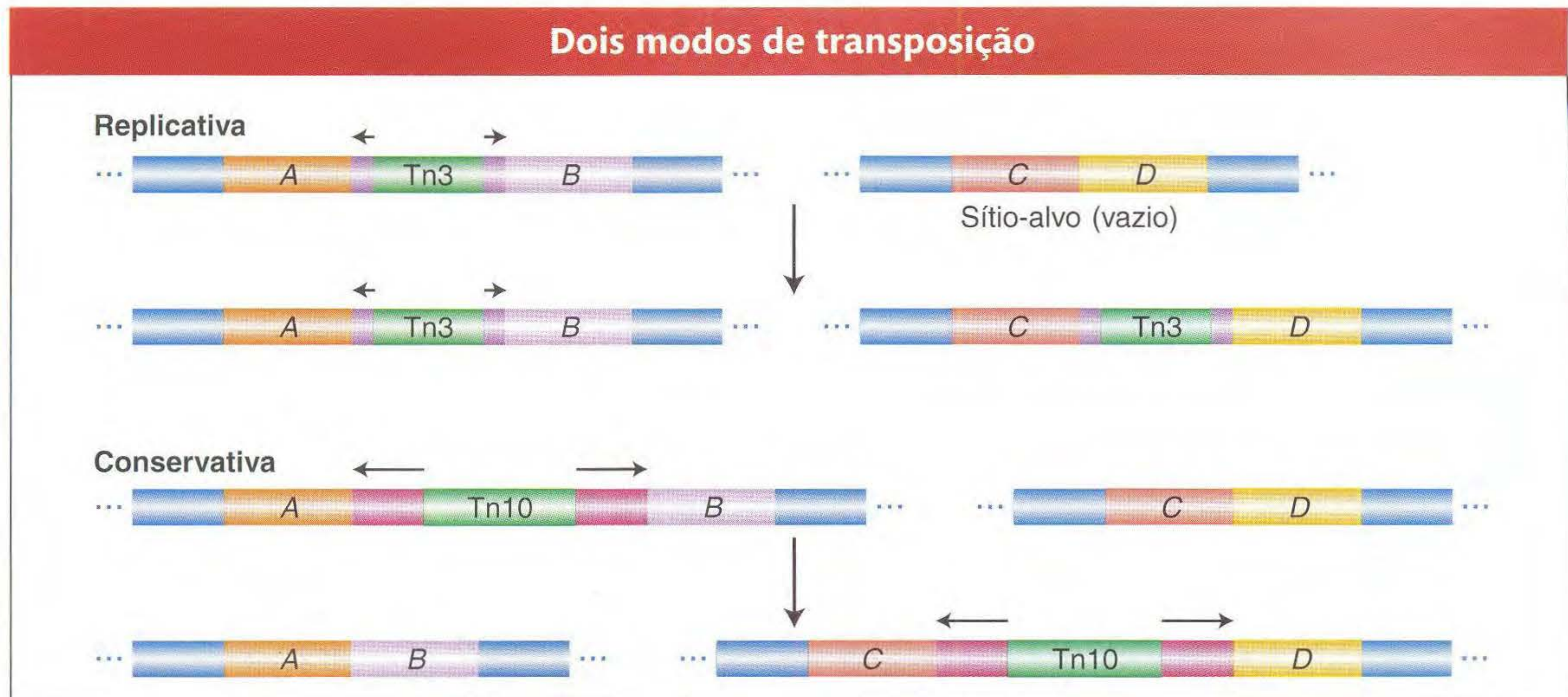


FIG. 14.9 A transposição de elemento móvel pode ser replicativa ou conservativa. Veja o texto para detalhes. [Adaptada com permissão de *Nature Reviews: Genetics* 1, n. 2, p. 138, Figure 3, November 2000, "Mobile Elements and the Human Genome," E.T. Luning Prak and H.H. Kazazian, Jr. Copyright 2000 by Macmillan Magazines Ltd.]

Transposição replicativa Como esse mecanismo é um pouco complicado, ele será descrito aqui em mais detalhes. Como a Figura 14.9 ilustra, uma cópia de Tn3 é produzida de uma única cópia inicial, produzindo duas cópias de Tn3.

A Figura 14.10 mostra os detalhes dos intermediários na transposição de Tn3 de um plasmídeo (o doador) para outro plasmídeo (o alvo). Durante a transposição, os plasmídeos doador e receptor são temporariamente fundidos para formar um plasmídeo duplo. A formação desse intermediário é catalisada pela transposase codificada por Tn3, que faz cortes unifilamentares nas duas pontas de Tn3 e cortes desencontrados na sequência-alvo (lembre dessa reação na Figura 14.8) e junta as pontas livres, formando um círculo fusionado chamado de **cointegrado**. O elemento transponível é duplicado no evento de fusão. O cointegrado, então, resolve-se por um evento tipo recombinação que transforma o cointegrado em dois círculos menores, deixando uma cópia do elemento de transposição em cada plasmídeo. O resultado é que uma cópia permanece no local original do elemento, enquanto a outra é integrada em uma nova posição genômica.

Transposição conservativa Alguns transposons, tais como Tn10, são removidos do cromossomo e integrados ao DNA-alvo. Nesses casos, a replicação do DNA do elemento não ocorre, e o elemento é perdido do sítio original do cromossomo (veja Figura 14.9). Como a transposição replicativa, essa reação é iniciada pela transposase do elemento codificado, que corta as pontas do transposon. Entretanto, em contraste com a transposição replicativa, a transposase retira o elemento do sítio doador. Ela então faz um corte desencontrado e insere o elemento no sítio-alvo. Voltaremos a tratar desse mecanismo, com mais detalhes, em uma discussão da transposição de elementos transponíveis eucarióticos, incluindo a família *Ac/Ds* de milho.

Nos procariontes, a transposição ocorre por, pelo menos, duas vias diferentes. Alguns elementos de transposição podem replicar uma cópia do elemento em um sítio-alvo, deixando para trás uma cópia no sítio original. No outro caso, a transposição consiste na remoção direta do elemento e sua reinserção em um novo sítio.

► 14.3 Elementos de Transposição em Eucariontes

Embora os elementos de transposição tenham sido primeiro descobertos em milho, os primeiros elementos eucarióticos a serem caracterizados molecularmente foram isolados de leveduras mutantes e genes de *Drosophila*. Os elementos transponíveis eucarióticos enquadram-se em duas classes: classe 1 de retrotransposons e DNA classe 2 de transposons. Primeira classe a ser isolada, os retrotransposons não eram como os elementos IS procarióticos e transposons.

Classe I: retrotransposons

O laboratório de Gerry Fink está entre os primeiros a usar leveduras como um organismo-modelo para estudar a regulação gênica eucariótica. Ao longo dos anos, ele e seus colegas isolaram milhares de mutações no gene *HIS4*, que codifica uma das enzimas na via que leva à síntese do aminoácido histidina.

Eles isolaram mais de 1.500 mutantes espontâneos *HIS4* e verificaram que 2 deles tinham um fenótipo mutante instável. Os mutantes instáveis eram mais de 1.000 vezes capazes de reverter ao tipo selvagem que os outros mutantes *HIS4*. Simbolicamente, dizemos que esses mutantes instáveis reverteram de his^- para His^+ (letras maiúsculas e sinais de expoentes são usados para indicar tipo selvagem, enquanto letras minúsculas e um sinal menos de expoente ou número da

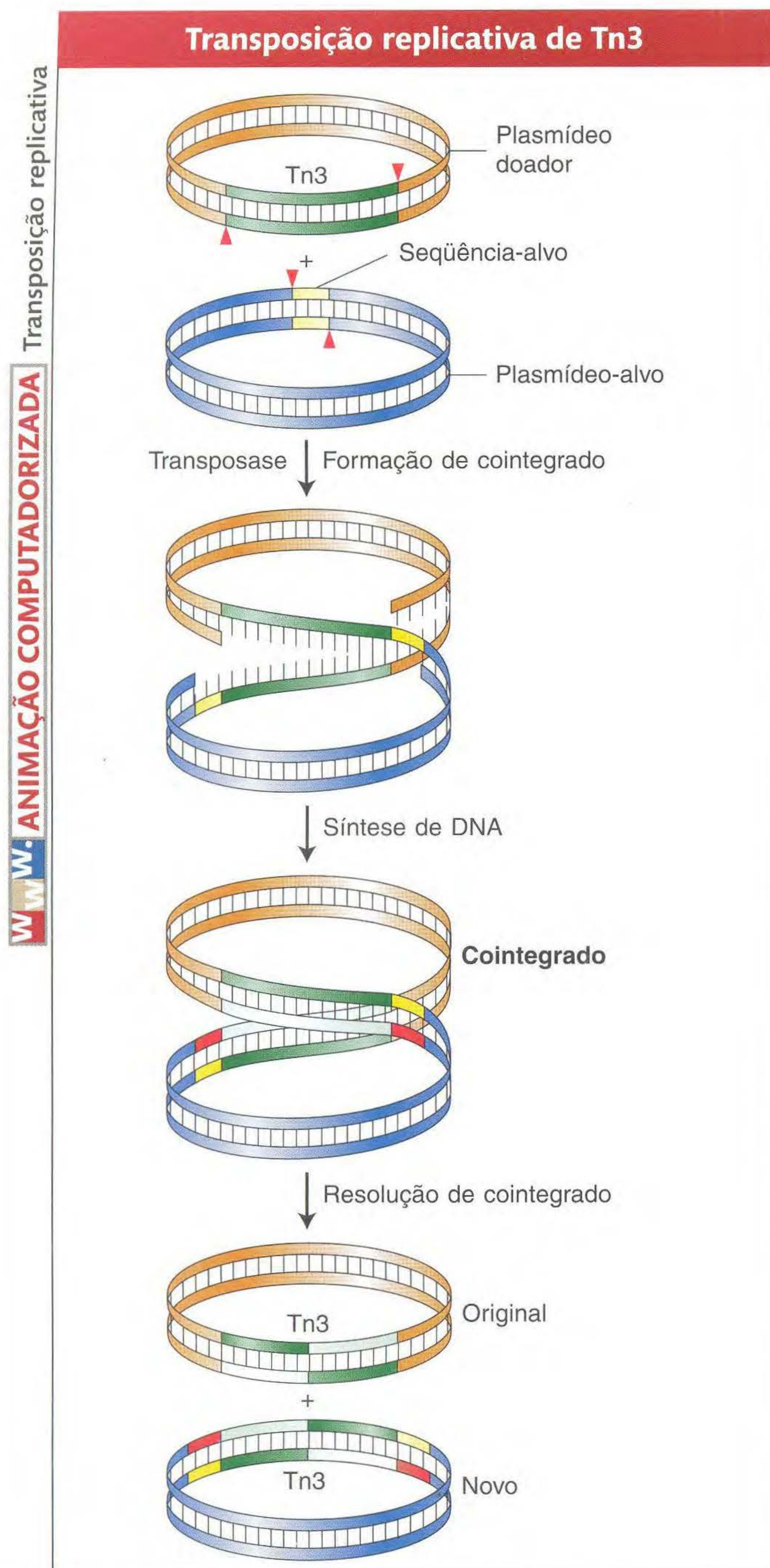


FIG. 14.10 Transposição replicativa de Tn3 ocorre por um intermediário cointegrado. [De Robert J. Brooker, *Genetics: Analysis and Principles*, Fig. 18-14. Benjamin-Cummings, 1999.]

mutação indicam um mutante). Como os mutantes *E. coli gal⁻*, essas leveduras mutantes foram encontradas abrigando uma grande inserção de DNA no gene *HIS4*. A inserção ficou muito similar àquela de um grupo de elementos de transposição já caracterizados em leveduras, chamados de **elementos Ty**. Existem, de fato, cerca de 35 cópias do elemento inserido, chamadas de *Ty1*, no genoma de levedura.

A clonagem dos elementos desses alelos mutantes levou à descoberta surpreendente de que as inserções não pare-

ciam os elementos bacterianos IS ou transposons. Ao contrário, elas se assemelham a uma classe bem caracterizada de vírus animais chamados de retrovírus. Um **retrovírus** é um vírus de RNA unifilamentar que emprega um DNA bifilamentar intermediário para replicação. O RNA é copiado em DNA pela enzima **transcriptase reversa**. O DNA bifilamentar é integrado aos cromossomos hospedeiros, dos quais é transcrito para produzir genoma de RNA viral e proteínas que formam novas partículas virais. Quando integrado aos cromossomos hospedeiros como DNA bifilamentar, a cópia bifilamentar de DNA do genoma retroviral é chamada de um **provírus**. O ciclo de vida de um retrovírus típico é mostrado na Figura 14.11. Alguns retrovírus, tais como o vírus do tumor mamário de camundongo (MMTV) e o vírus do sarcoma de Rous (RSV), são responsáveis pela indução de tumores cancerosos.

A Figura 14.12 mostra a similaridade de estrutura e o conteúdo gênico de um retrovírus e o elemento *Ty1* isolado dos mutantes *HIS4*. Ambos são flanqueados por sequências de **longa repetição terminal (LTR)** que têm centenas de pares de bases de tamanho. Ambos contêm os genes *gag* e *pol*.

Os retrovírus codificam pelo menos três proteínas que tomam parte na replicação viral: os produtos dos genes *gag*, *pol* e *env*. A proteína codificada por *gag* tem um papel na maturação do genoma de RNA, *pol* codifica a importante transcriptase reversa e *env* codifica a proteína estrutural que circunda o vírus. Essa proteína é necessária para que o vírus saia da célula para infectar outras células. Curiosamente, os elementos *Ty1* têm genes relacionados a *gag* e *pol*, mas não a *env*. Essas características levam à hipótese de que, como os retrovírus, os elementos *Ty1* são transcritos em RNA que são copiados em DNA bifilamentar pela transcriptase reversa. Entretanto, ao contrário dos retrovírus, os elementos *Ty1* não podem deixar a célula, pois não codificam *env*. Em vez disso, as cópias de DNA bifilamentar são inseridas de volta no genoma da mesma célula. Essas etapas são diagramadas na Figura 14.13.

Em 1985, David Garfinkel, Jef Boeke e Gerald Fink mostraram que, como os retrovírus, os elementos *Ty* de fato são transpostos por um RNA intermediário. A Figura 14.14 diagrama o esquema experimental. Eles começaram alterando um elemento *Ty1* de levedura, clonado em um plasmídeo. Primeiro, próximo a uma extremidade de um elemento, eles inseriram um promotor que pode ser ativado pela adição de galactose ao meio. Segundo, eles introduziram um íntron de outro gene de levedura na região codificante do transposon *Ty*.

A adição de galactose aumenta muito a frequência de transposição do elemento *Ty* alterado. Essa frequência aumentada sugere a participação de RNA, porque a galactose estimula a transcrição do DNA *Ty* em RNA, começando em um promotor sensível a galactose. O resultado experimental chave, entretanto, é o destino do DNA *Ty* transposto. Os pesquisadores encontraram que o íntron tinha sido removido do DNA *Ty* transposto. Como os íntrons são recompostos apenas no curso do processamento de RNA (veja o Capítulo 8), o DNA *Ty* transposto deve ter sido copiado de um RNA

Ciclo de vida de um retrovírus

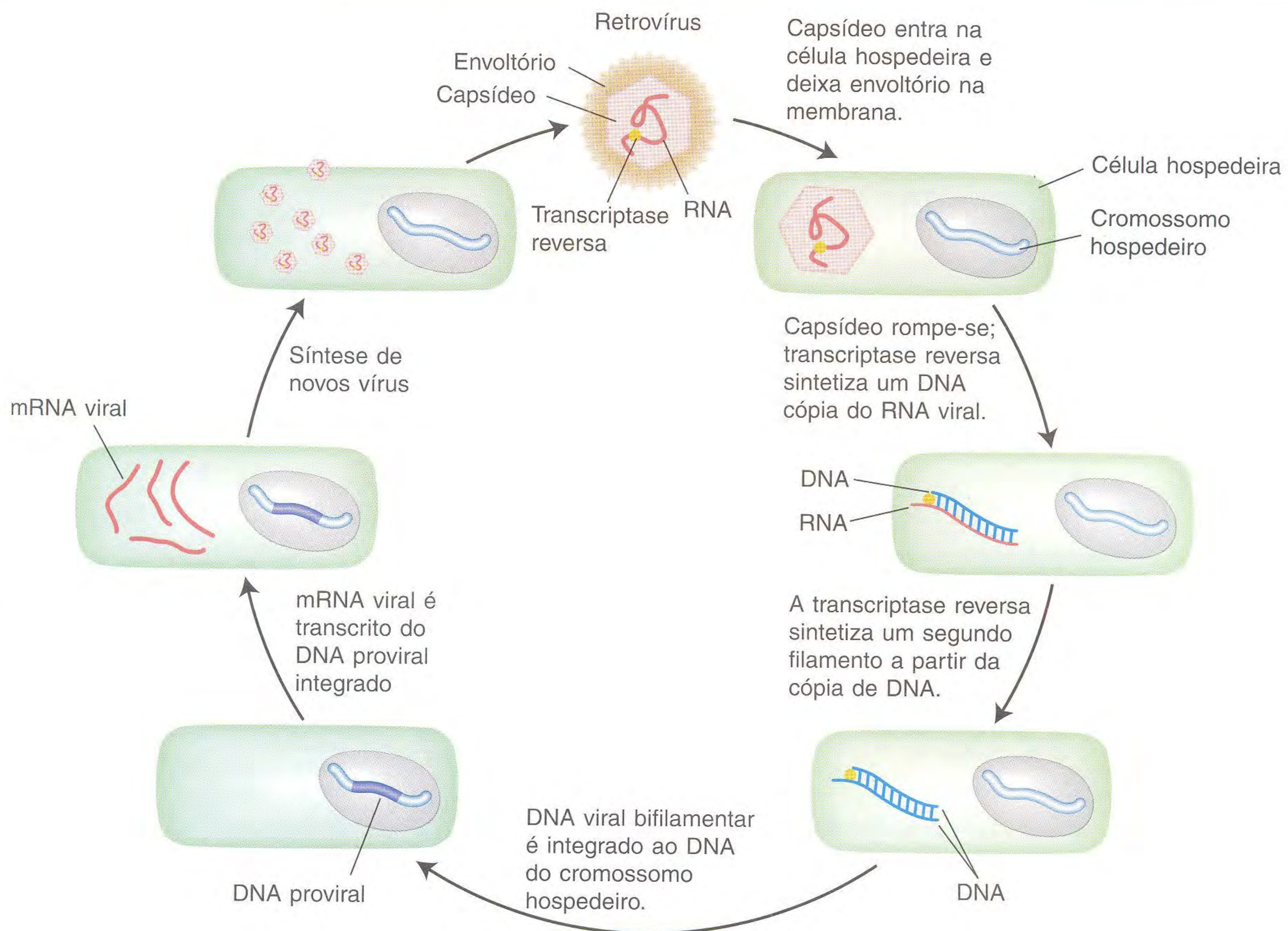


FIG. 14.11 O genoma de RNA do retrovírus sofre transcrição reversa em DNA bifilamentar dentro da célula hospedeira.

Retrotransposons têm características em comum com retrovírus

(a) Um retrovírus, MoMLV



(b) Ty1 em levedura



(c) Cópia em Drosophila



(d) L1, um LINE humano

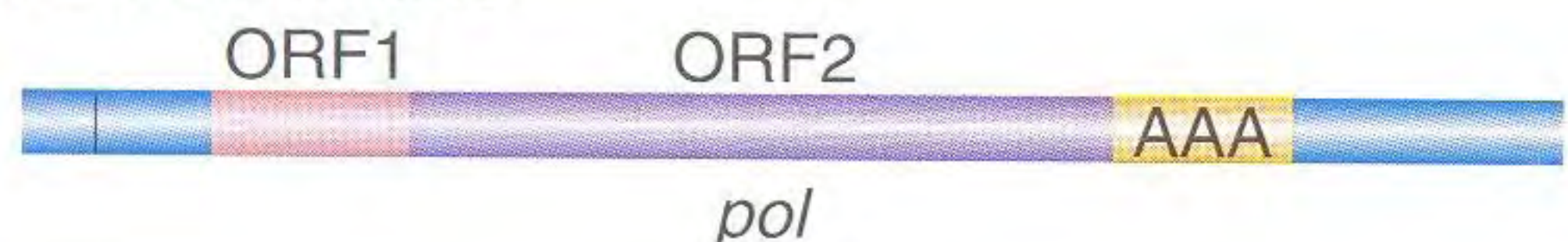


FIG. 14.12 Comparação estrutural de um retrovírus com retrotransposons encontrados em genomas eucarióticos. (a) Um retrovírus, vírus de leucemia murina Moloney (MoMLV), de camundongo. (b) Um retrotransposon, Ty1, em levedura. (c) Um retrotransposon, copia, em Drosophila. (d) Um longo elemento intercalar (LINE) em humanos. Abreviações: LTR, longa repetição terminal; ORF, matriz de leitura aberta.

Um retrotransposon é transposto por um RNA intermediário

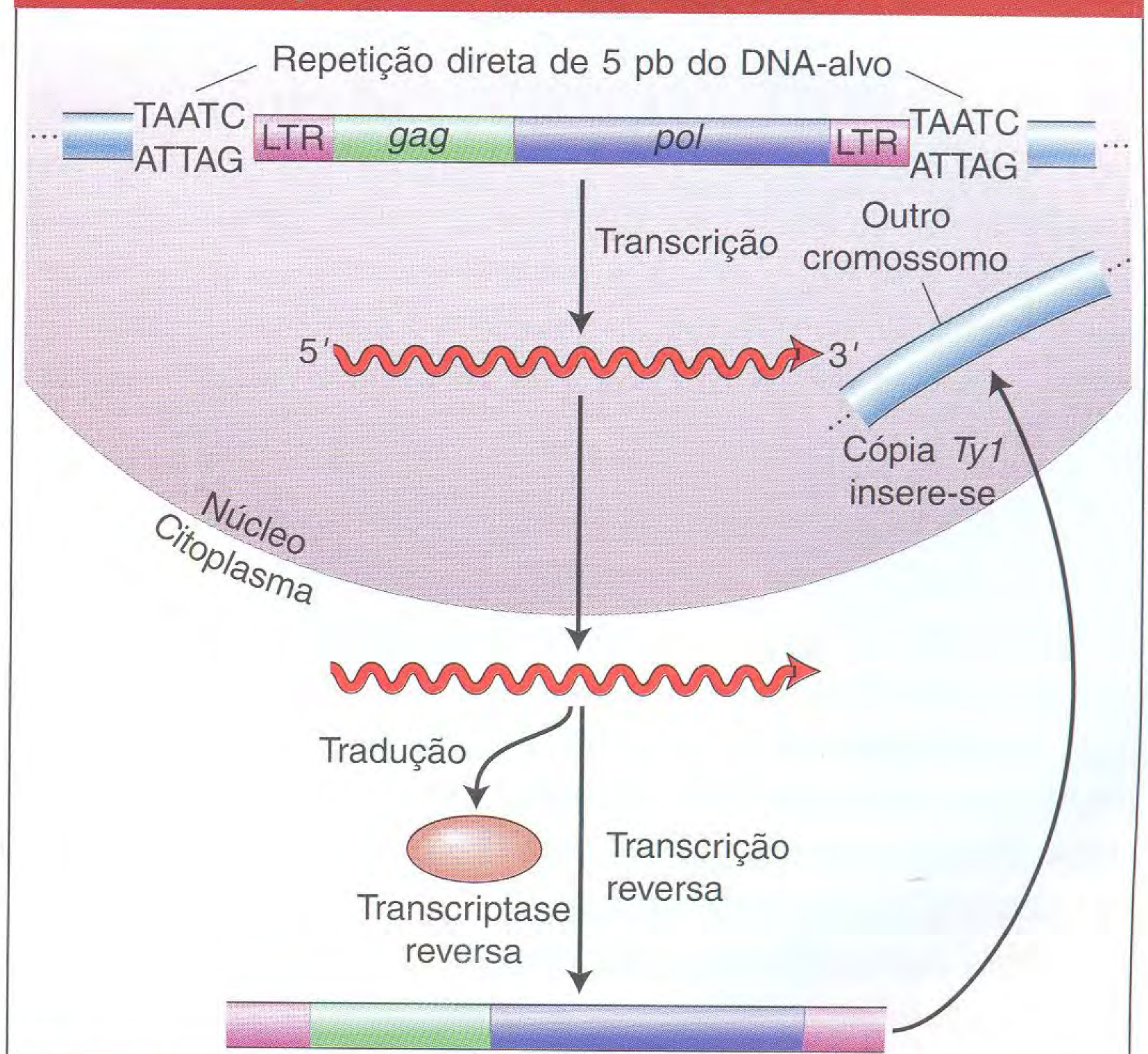


FIG. 14.13 Um RNA transcrito de retrotransposon sofre transcrição reversa em DNA, por uma transcriptase reversa codificada pelo retrotransposon. Uma cópia de DNA é inserida em um novo local no genoma.

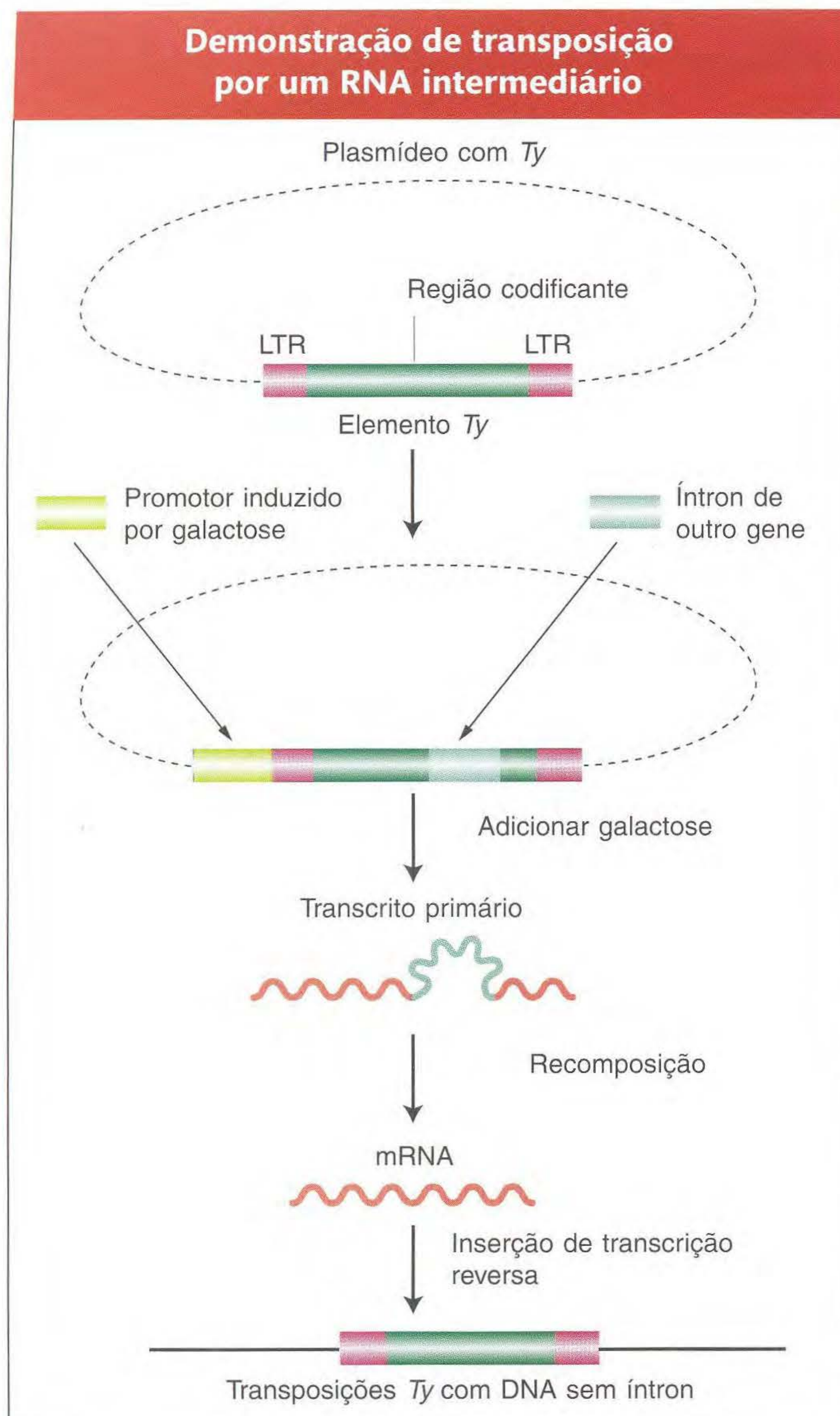


FIG. 14.14 Um elemento Ty é alterado pela adição de um íntron e um promotor que pode ser ativado pela adição de galactose. As seqüências íntron são recompostas antes da transcrição reversa. [De H. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, and J. Darnell, *Molecular Cell Biology*, 3rd ed., p. 332. Copyright 1995 by Scientific American Books.]

intermediário. A conclusão foi que o RNA é transcrito do elemento Ty original e recomposto. O mRNA recomposto sofre transcrição reversa de volta ao DNA bifilamentar, que é então integrado ao cromossomo de levedura. Os elementos de transposição que empregam a transcriptase reversa para transpor por meio de um RNA intermediário são chamados de **retrotransposons**. Eles são também conhecidos como **elementos de transposição classe 1**. Os retrotransposons tais como Ty1 que têm *longas repetições terminais* em suas pontas são chamados de **retrotransposons LTR**.

Várias mutações espontâneas isoladas durante os anos em *Drosophila* também mostraram conter inserções de retrotransposons. Os **elementos tipo cópia** de *Drosophila* são estruturalmente similares aos elementos Ty1 e aparecem em 10 a 100 posições no genoma de *Drosophila* (veja Figura

14.12c). Algumas mutações clássicas de *Drosophila* resultam da inserção de tipo cópia de outros elementos. Por exemplo, a mutação *white-apricot* (w^a) para cor de olho é causada pela inserção de um elemento da família cópia no locus *white*. Também foi mostrado que a inserção de retrotransposons LTR em genes de plantas (inclusive milho) contribui para mutações espontâneas nesse reino.

Os elementos de transposição que se transpõem por meio de intermediários de RNA predominam em eucariontes. Os retrotransposons, também conhecidos como elementos classe 1, codificam uma transcriptase reversa que produz uma cópia de DNA bifilamentar (a partir de um RNA intermediário) que é capaz de se integrar em uma nova posição no genoma.

Transposons de DNA

Alguns elementos móveis encontrados em eucariontes parecem transpor-se por mecanismos similares aos de bactéria. Como ilustrado na Figura 14.9 para os elementos IS e transposons, o que se insere em uma nova posição no genoma é ou o próprio elemento, ou uma cópia do elemento. Os elementos que se transpõem dessa maneira são classificados como **elementos classe 2**, ou **transposons de DNA**. Os primeiros elementos transponíveis descobertos por McClintock em milho hoje são conhecidos como transposons de DNA. Entretanto, os primeiros transposons de DNA a serem molecularmente caracterizados foram os elementos *P* em *Drosophila*.

Elementos P De todos os elementos de transposição em *Drosophila*, os mais curiosos e úteis para o geneticista são os **elementos P**. Esses elementos foram descobertos por Margaret Kidwell, que estava estudando a **disgenesia híbrida**, um fenômeno que ocorre quando as fêmeas de linhagens de laboratório de *D. melanogaster* são cruzadas com machos derivados de populações naturais. Em tais cruzamentos, diz-se que os estoques de laboratório possuem um **citotipo M** (tipo celular), e os estoques naturais, um **citotipo P**. Em um cruzamento de M (fêmea) × P (macho), a prole apresenta uma gama de fenótipos surpreendentes que são manifestos na linhagem germinativa, incluindo a esterilidade, uma alta taxa de mutação e uma alta frequência de anomalias cromossômicas e não-disjunção (Figura 14.15). Essa prole híbrida é **disgênica**, ou biologicamente deficiente (daí a expressão **disgenesia híbrida**). Curiosamente, o cruzamento recíproco, P (fêmea) × M (macho), não produz prole disgênica. Uma observação importante é que uma grande percentagem de mutações disgenicamente induzidas são instáveis; isto é, elas reverterem ao tipo selvagem ou a outros alelos mutantes em frequências muito altas. Essa instabilidade geralmente é restrita à linhagem germinativa de uma mosca que possua o citotipo M.

Os mutantes instáveis de *Drosophila* têm similaridades aos mutantes instáveis de milho caracterizados por McClintock. Os pesquisadores formularam a hipótese de que as mutações disgênicas são causadas pela inserção de elementos de trans-

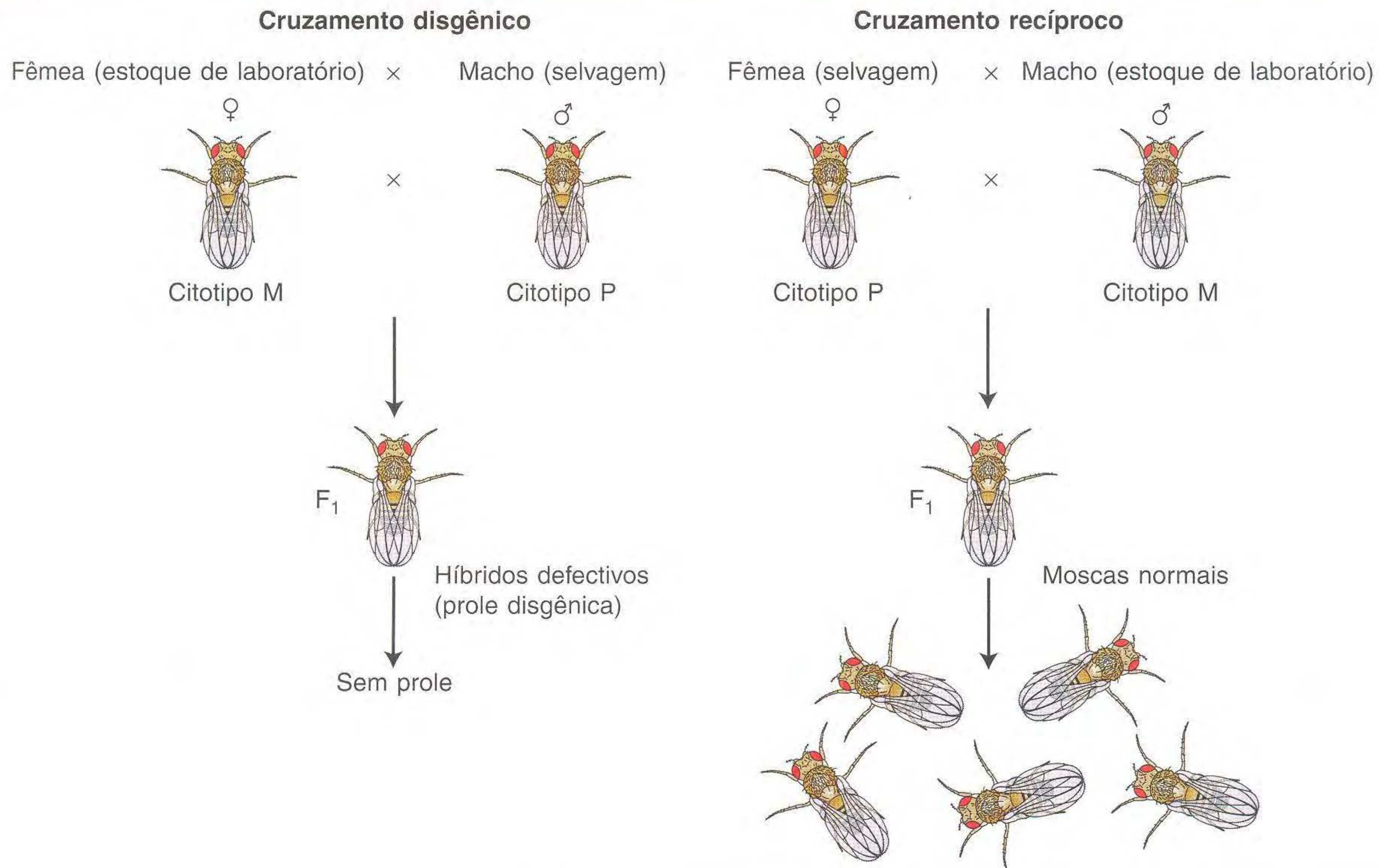
Disgenesia híbrida em *Drosophila*

FIG. 14.15 Na disgenesia híbrida, um cruzamento entre uma fêmea do estoque de laboratório e um macho tipo selvagem produz uma prole defectiva. Veja texto para detalhes.

posição em genes específicos, tornando-os assim inativos. De acordo com essa visão, a reversão geralmente resulta da excisão dessas seqüências inseridas. Essa hipótese foi testada criticamente isolando mutações disgênicas instáveis no locus da cor de olho *white*. A maioria das mutações foram causadas pela inserção de um elemento de transposição no gene *white*⁺. O elemento, chamado de *elemento P*, foi encontrado presente em 30 a 50 cópias por genoma nas linhagens P, mas estava totalmente ausente nas linhagens M. Os elementos P variam de tamanho, indo de 0,5 a 2,9 kb de comprimento. Essa diferença de tamanho é devida à presença de muitos elementos P defectivos dos quais a metade do elemento foi deletada. O elemento P de tamanho total lembra os transposons simples de bactérias, pois suas pontas são curtas (31 pb) repetições

invertidas, e ele codifica uma transposase. Entretanto, esse gene de transposase eucariótica contém três íntrons e quatro éxons (Figura 14.16).

Por que os elementos P não causam problemas nas linhagens P? As linhagens P foram propostas contendo elementos P e um repressor que impede a transposição dos elementos P dentro do genoma. De acordo com esse modelo, que é mostrado na Figura 14.17, os elementos P, como os elementos bacterianos IS e Tn, codificam uma transposase que é responsável por sua mobilização. Além disso, os elementos P codificam um repressor cujo papel é evitar a produção de transposase, bloqueando assim a transposição. Por algum motivo, a maioria das linhagens laboratoriais não têm elementos P e, conseqüentemente, nenhum repressor no citoplasma. Nos híbridos do cruzamento M (fêmea, sem elementos P) × P (macho, elementos P), os elementos P no zigoto recém-formado estão em um ambiente sem repressor porque o espermatozóide contribui com seu genoma (com elementos P), mas nenhum citoplasma (com o repressor). Os elementos P derivados do genoma masculino agora podem transpor-se pelo genoma diplóide, causando uma variedade de danos à medida que se inserem nos genes que causam mutações. Esses eventos moleculares são expressos como as várias manifestações de disgenesia híbrida. Por outro lado, o cruzamento P (fêmea) × M (macho) não resulta em disgenesia, pois, nesse caso, o citoplasma do zigoto contém o repressor P.

Uma questão curiosa permanece sem resposta: Por que as linhagens de laboratório não têm elementos P, enquanto as

Estrutura do elemento P

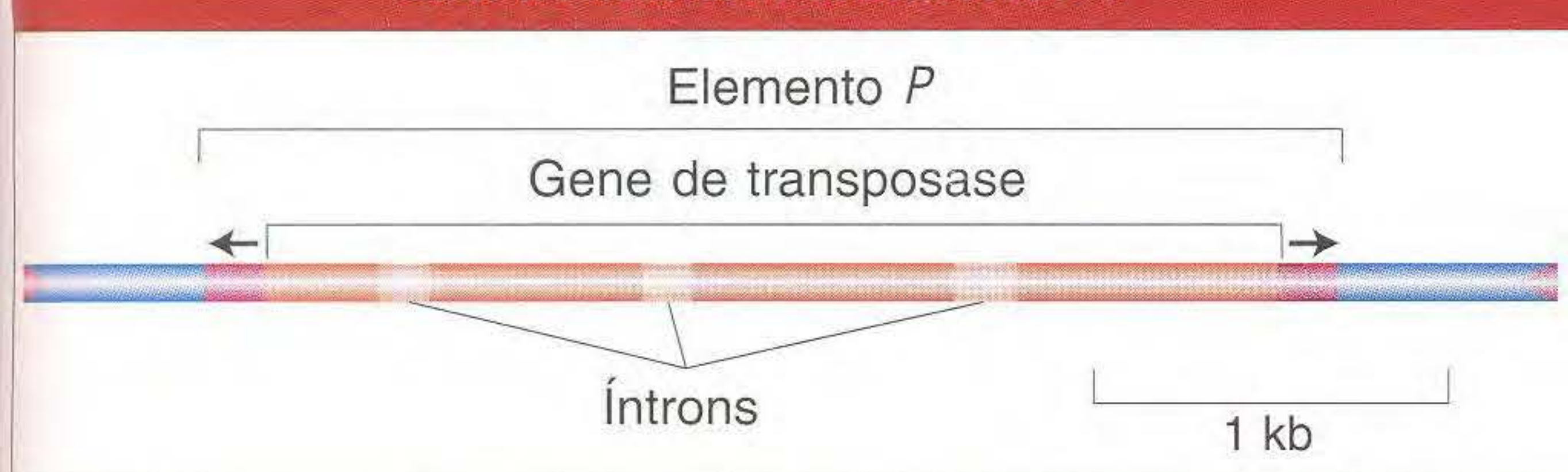


FIG. 14.16 Análise de seqüência de DNA do elemento P de 2,9 kb revela um gene, composto de quatro éxons e três íntrons, que codifica transposase. Uma perfeita repetição invertida de 31 pb reside em cada ponta do elemento. [De G. Robin, in J.A. Shapiro, Ed., *Mobile Genetic Elements*, pp. 329-361. Academic Press, 1983.]

A inserção de elementos *P* causa mutações subjacentes à disgenesia híbrida

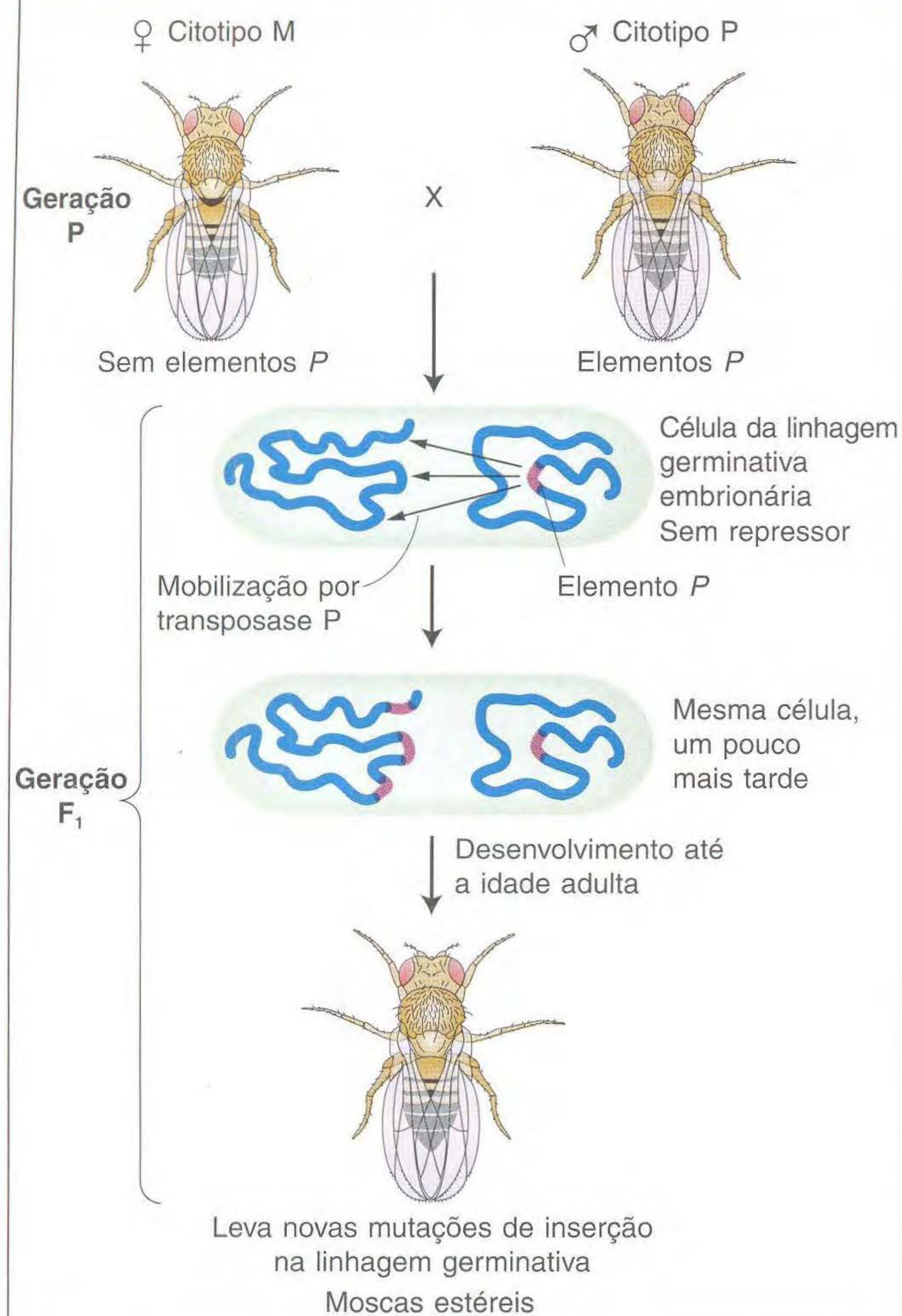


FIG. 14.17 Eventos moleculares subjacentes a disgenesia híbrida. Os cruzamentos do macho de *Drosophila* portador de transposase *P* com uma fêmea que não tenha elementos *P* funcionais produz mutações na linhagem germinativa da prole F_1 causadas pelas inserções do elemento *P*. Os elementos *P* são capazes de mover-se causando mutações, pois os espermatozoides do macho não trazem repressores consigo.

linhagens tipo selvagem têm elementos *P*? Uma hipótese é que as linhagens de laboratório atuais descendem de isolados originais colhidos da natureza por Morgan e seus estudantes há quase um século. Em algum ponto entre a captura dessas linhagens originais e as presentes, os elementos *P* se espalharam pela população natural, mas não pelas linhagens de laboratório. Essa diferença não foi percebida até que as linhagens selvagens fossem novamente capturadas e cruzadas com as linhagens de laboratório.

Embora o cenário exato de como os elementos *P* se espalharam pelas populações selvagens não esteja claro, o que está claro é que os elementos de transposição podem se espalhar rapidamente de alguns membros de uma população. Quanto a isso, a dispersão de elementos *P* assemelha-se à dispersão de transposons, levando genes de resistência para populações bacterianas antes suscetíveis.

A transposase ativadora catalisa a excisão e a integração

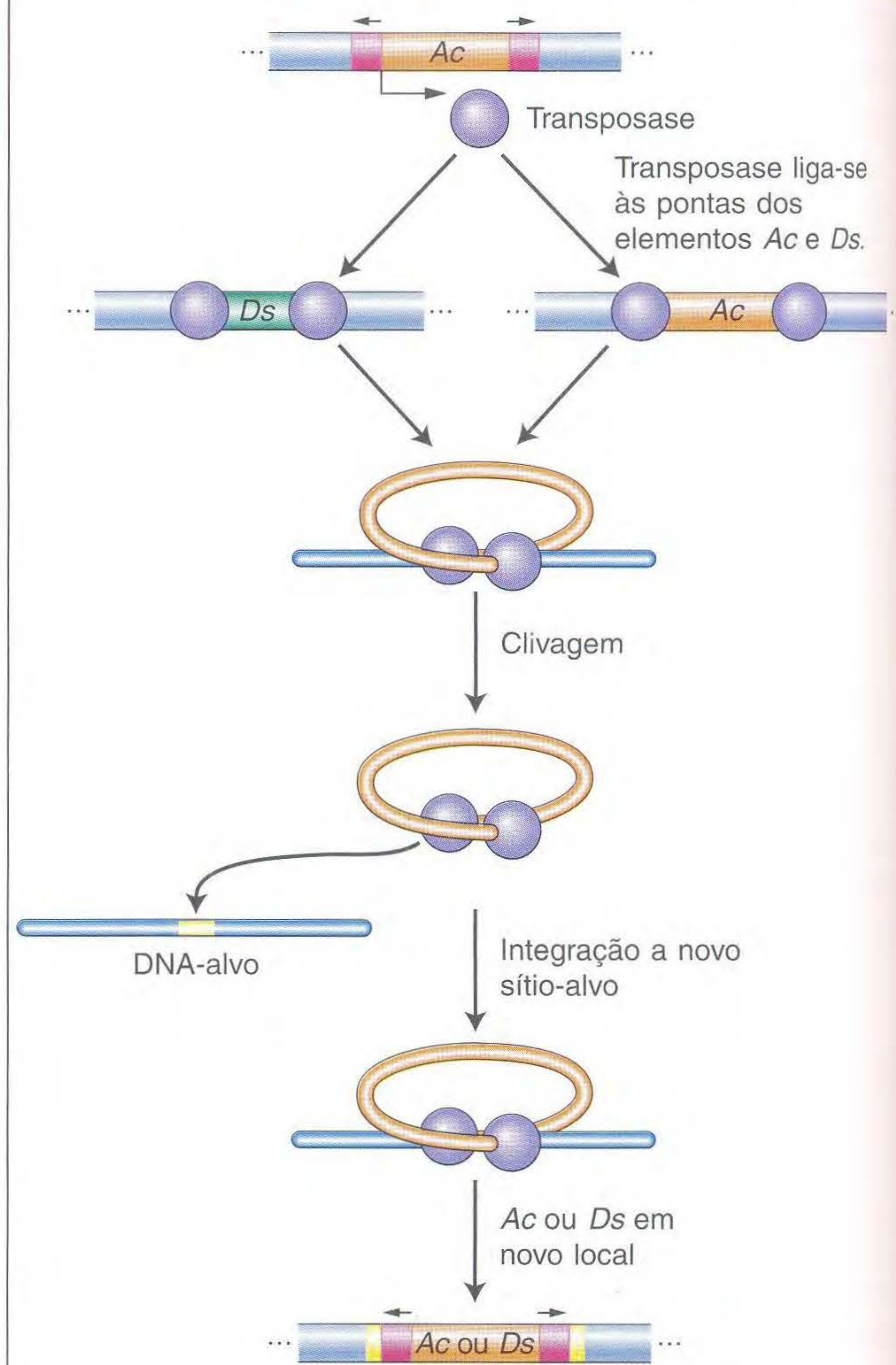


FIG. 14.18 Um elemento *Ac* em milho codifica uma transposase que liga suas próprias pontas ou as de um elemento *Ds*, removendo o elemento, clivando o sítio-alvo e permitindo que o elemento seja inserido em outra parte do genoma.

Voltando aos elementos de transposição em milho

Embora o agente causal responsável por mutantes instáveis primeiro tenha sido mostrado geneticamente como sendo os elementos de transposição em milho, há quase 50 anos antes os elementos *Ac* e *Ds* de milho foram isolados e mostrados como relacionados a transposons de DNA em bactérias e em outros eucariontes. Como o elemento *P* de *Drosophila*, *Ac* tem repetições terminais invertidas e codifica uma única proteína, a transposase. O elemento *Ds* não-autônomo não codifica transposase e, portanto, não pode transpor-se. Quando *Ac* está no genoma, sua transposase pode se ligar a ambas as extremidades dos elementos *Ac* e *Ds* e promover sua transposição (Figura 14.18).

Como notado antes neste capítulo, *Ac* e *Ds* são membros de uma única família de transposons, e existem outras famílias de elementos de transposição em milho. Cada família contém um elemento autônomo que codifica uma transposase

capaz de mover os elementos na mesma família, mas não mover elementos em outras famílias, pois a transposase só pode ligar-se às extremidades dos membros da família.

Embora alguns organismos, tais como leveduras, não tenham transposons de DNA, foram isolados elementos estruturalmente similares aos elementos *P* e *Ac* de muitas espécies de plantas e animais. De fato, a mutação do gene de pigmento responsável pelo mutante floral boca-de-leão, mostrado na Figura 14.5, é causada pela inserção de um elemento chamado *Tam3*, que é muito similar a *Ac*. Várias cópias de *Tam3* residem no genoma de boca-de-leão.

Os primeiros elementos de transposição conhecidos em milho são transposons de DNA que, estruturalmente, assemelham-se a transposons de DNA em bactérias e outros eucariontes. Os transposons de DNA codificam a transposase que corta o transposon do cromossomo e catalisa sua reinserção em outros locais cromossômicos.

Utilidade dos transposons de DNA para a descoberta do gene

Além de seu interesse como um fenômeno genético, os transposons de DNA tornaram-se ferramentas importantes para os geneticistas trabalharem com uma variedade de organismos. Sua mobilidade foi explorada para marcar genes para clonagem e inserir transgenes. O elemento *P* em *Drosophila* fornece um dos melhores exemplos de como os geneticistas exploram as propriedades dos elementos de transposição em eucariontes.

Uso de elementos *P* para marcar genes para clonagem

Os elementos *P* podem ser usados para criar mutações por inserção, marcar a posição dos genes e facilitar a clonagem dos genes. Os elementos *P* inseridos nos genes *in vivo* perturbam os genes aleatoriamente, criando mutantes com fenótipos diferentes. As moscas-das-frutas com fenótipos mutantes interessantes podem ser selecionadas para clonagem do gene mutante, que é marcado pela presença do elemento *P*, um método chamado de **marcação de transposon**. Após o gene interrompido ter sido clonado, os fragmentos do gene mutante podem ser usados como uma sonda para isolar o gene tipo selvagem.

Uso de elementos *P* para inserir genes Gerald Rubin e Allan Spradling mostraram que o elemento *P* de DNA pode ser usado como um veículo efetivo para transferir genes doadores para a linhagem germinativa de uma mosca receptora. Eles criaram o seguinte procedimento experimental (Figura 14.19). Suponha que a meta seja transferir o alelo *ry*⁺, que confere uma cor de olho característica, para o genoma da mosca. O genótipo receptor é homozigoto para a mutação *rosy* (*ry*⁻). Dessa linhagem são coletados embriões ao término de cerca de nove divisões nucleares. Nesse estágio, o embrião é uma célula multinucleada, e os núcleos destinados a formar células germinativas são agrupados em uma extre-

midade. (Os elementos *P* mobilizam apenas nas células da linhagem germinativa.) Dois tipos de DNA são injetados em embriões desse tipo. O primeiro é um plasmídeo bacteriano levando um elemento *P* defectivo no qual o gene *ry*⁺ foi inserido. O elemento *P* defectivo assemelha-se ao elemento *Ds* de milho por não codificar a transposase, mas ainda tem as pontas que ligam a transposase e permitem a transposição. Esse elemento deletado não é capaz de se transpor, e, assim, como foi mencionado antes, o plasmídeo auxiliar codificando a transposase também é injetado. As moscas que se desenvolvem desses embriões são ainda fenotipicamente mutantes *rosy*, mas sua prole inclui uma grande proporção de moscas *ry*⁺. A hibridização *in situ* confirmou que o gene *ry*⁺, juntamente com o elemento *P* deletado, foi inserido em um de vários locais cromossômicos distintos. Nenhum apareceu exatamente no locus normal do gene *rosy*. Esses novos genes *ry*⁺ são encontrados como herdados de modo mendeliano estável.

Como o elemento *P* pode se transpor em *Drosophila*, essas aplicações são restritas em seu uso. Em contraste, o elemento *Ac* de milho é capaz de se transpor após sua introdução nos genomas de espécies de plantas, incluindo a mostarda *Arabidopsis*, alface, cenoura, arroz, cevada, dentre muitas outras. Como os elementos *P*, *Ac* foi criado pelos geneticistas para uso no isolamento de genes por marcação de transposon. Desse modo, *Ac*, o primeiro elemento de transposição descoberto por Barbara McClintock, serviu como uma ferramenta importante da genética de plantas mais de 50 anos depois.

Os transposons de DNA foram modificados e usados por cientistas de dois modos importantes: (1) para fazer mutantes que podem ser identificados molecularmente pela presença de uma marcação de transposon e (2) como vetores que podem introduzir genes exógenos em um cromossomo.

► 14.4 O Genoma Dinâmico: Mais Elementos de Transposição do que se Imaginava

Como você já viu, os elementos de transposição foram primeiro descobertos com o uso de enfoques genéticos. Nesses estudos, os elementos tornaram sua presença conhecida quando foram transpostos para um gene ou em sítios de quebra cromossômica ou rearranjo. Após o DNA de elementos de transposição ser isolado de mutações instáveis, os cientistas puderam usar esse DNA como sondas moleculares para determinar se havia mais cópias correlatas no genoma. Em todos os casos, pelo menos várias cópias do elemento sempre estavam presentes no genoma e, em alguns casos, até muitas centenas.

Os cientistas pensaram sobre a prevalência dos elementos de transposição nos genomas. Haveria outros elementos de transposição no genoma que permaneceram desconhecidos porque não causaram uma mutação que poderia ser estudada no laboratório? Existiriam outros elementos de transposição

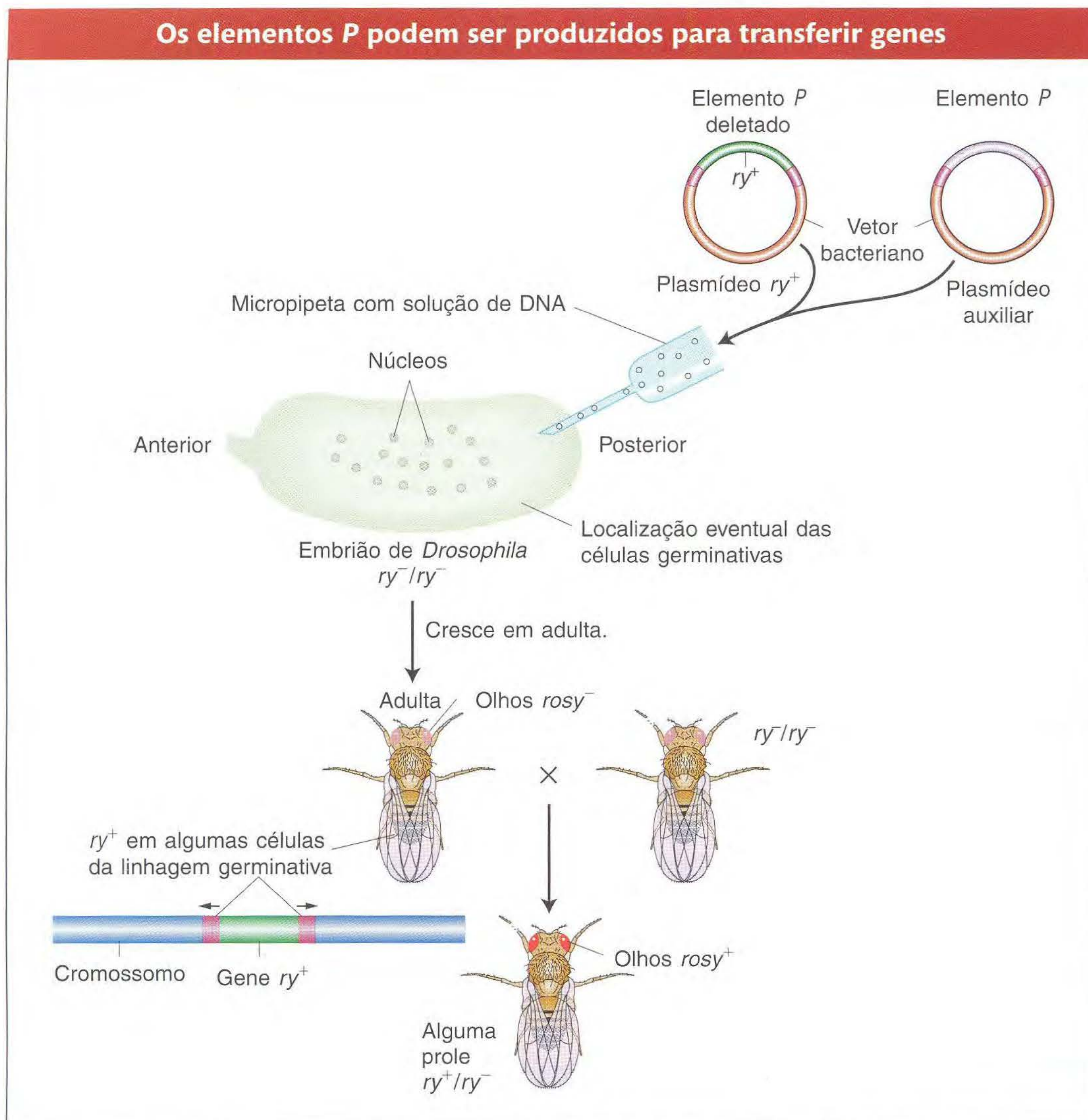


FIG. 14.19 Transferência gênica mediada por elemento P em *Drosophila*. O gene de cor de olho *rosy*⁺ (*ry*⁺) é inserido em um elemento P deletado levado em um vetor bacteriano. Ao mesmo tempo, é usado um plasmídeo auxiliar portador de um elemento P intacto. Ambos são injetados em um embrião *ry*⁻, onde *ry*⁺ transpõe-se com o elemento P para os cromossomos da linhagem de células germinativas.

na grande maioria dos organismos que não eram passíveis de análise genética? Perguntado de outro modo, organismos sem mutações induzidas por elementos de transposição teriam elementos de transposição em seus genomas? Essas perguntas são remanescentes da pergunta: Se uma árvore cai na floresta, ela faria barulho se ninguém estiver escutando?

Genomas grandes são amplamente elementos de transposição

Muito antes do advento dos projetos de sequenciamento de DNA, os cientistas usando uma variedade de técnicas bioquímicas descobriram que o conteúdo de DNA (chamado de **valor C**) varia muito em eucariontes, e não está correlacionado à complexidade biológica. Por exemplo, os genomas de salamandras são 20 vezes maiores do que o genoma humano, enquanto o genoma de cevada é mais de 10 vezes maior que o genoma de arroz, uma gramínea correlata. A falta de correlação entre o tamanho do genoma

e a complexidade biológica de um organismo é conhecida como **paradoxo do valor C**.

Cevada e arroz são ambos cereais e, assim, seu conteúdo gênico deve ser similar. Entretanto, se os genes são um componente relativamente constante dos genomas de organismos multicelulares, o que é responsável pelo paradoxo do valor C? Com base nos resultados de experimentos adicionais, os cientistas foram capazes de determinar que as seqüências de DNA que são repetidas milhares e, mesmo, centenas de milhares de vezes constituem uma grande fração dos genomas eucarióticos, e que alguns genomas contêm muito mais DNA repetitivo do que outros.

Graças a muitos projetos recentes para sequenciar os genomas de uma ampla variedade de grupos (incluindo *Drosophila*, humanos, camundongo, *Arabidopsis* e arroz), hoje sabemos que existem muitas classes de seqüências repetitivas nos genomas de organismos superiores, e que alguns são similares aos transposons de DNA e retrotransposons




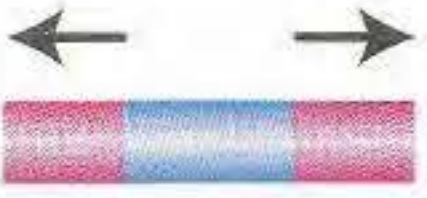
Tipos de elementos de transposição no genoma humano					
Elemento	Transposição	Estrutura	Tamanho	Número de cópias	Fração do genoma
LINE	Autônoma		1–5 kb	20.000–40.000	21%
SINE	Não-autônoma		100–300 pb	1.500.000	13%
Transposons de DNA	Autônoma		2–3 kb	300.000	3%
	Não-autônoma		80–3.000 pb		

FIG. 14.20 Várias classes gerais de elementos de transposição são encontradas no genoma humano. [Reimpresso com permissão de *Nature* 409, 880 (15 February, 2001). “Initial Sequencing and Analysis of the Human Genome”, The International Human Genome Sequencing Consortium. Copyright 2001 by Macmillan Magazines Ltd.]

mostrados como responsáveis por mutações em plantas, leveduras e insetos. Mais marcantemente, essas seqüências constituem a maioria do DNA nos genomas de eucariontes multicelulares.

Em vez de correlacionar com o conteúdo gênico, o tamanho do genoma freqüentemente está correlacionado com a quantidade de DNA no genoma que é derivado de elementos de transposição. Os organismos com genomas grandes têm muitas seqüências que se assemelham a elementos de transposição, enquanto organismos com genomas pequenos têm bem menos. Dois exemplos, um do genoma humano e outro de uma comparação dos genomas de gramíneas, ilustram esse ponto. As características estruturais dos elementos de transposição que são encontradas nos genomas humanos estão resumidas na Figura 14.20 e serão citadas na seção seguinte.

O paradoxo do valor C é a falta de correlação entre o tamanho do genoma e a complexidade biológica. Os genes constituem apenas uma pequena proporção dos genomas de organismos multicelulares. O tamanho do genoma corresponde à quantidade de seqüências de elementos de transposição, e não ao conteúdo gênico.

Elementos de transposição no genoma humano

Quase metade do genoma humano deriva-se de elementos de transposição. A grande maioria desses elementos de transposição são de dois tipos de retrotransposons, chamados de **elementos longos intercalados**, ou **LINE**, e **elementos curtos intercalados**, ou **SINE** (veja Figura 14.20). Os LINE movem-se como um retrotransposon com a ajuda de um elemento codificado por uma transcriptase reversa, mas não têm algumas características estruturais de elementos tipo retrovírus, incluindo LTR (veja Figura 14.12d). Os SINE podem ser mais bem descritos como LINE não-autônomos, porque têm as características estruturais de LINE, mas não codificam sua própria transcriptase reversa. Supostamente, eles são mobili-

zados por enzimas transcriptase reversa codificadas por LINE que residem no genoma.

O SINE mais abundante em humanos é chamado de **Alu**, porque contém um sítio-alvo para a enzima de restrição Alu. O genoma humano contém mais de 1 milhão de seqüências *Alu* inteiras e parciais, distribuídas entre genes e dentro de íntrons. Essas são seqüências *Alu* que constituem mais de 10% do genoma humano. Toda a seqüência *Alu* tem cerca de 200 nucleotídeos e possui uma lembrança marcante com o 7SL RNA, um RNA que é parte de um complexo pelo qual novos polipeptídeos sintetizados são secretados pelo retículo endoplasmático. Supostamente, as seqüências *Alu* originaram-se como transcritos reversos dessas moléculas de RNA.

Há cerca de 20 vezes mais DNA no genoma humano derivado de elementos de transposição do que DNA codificante de todas as proteínas humanas. A Figura 14.21 ilustra o número e a diversidade de elementos de transposição presentes no genoma humano, usando como exemplo as posições de *Alu* individuais, outros SINE e LINE na vizinhança de um gene humano típico.

O genoma humano parece ser típico de um organismo multicelular na abundância e distribuição de elementos de transposição. Assim, uma questão óbvia é: Como as plantas e animais sobrevivem com tantas inserções nos genes e tanto DNA móvel no genoma? Primeiro, com relação à função do gene, todos os elementos mostrados na Figura 14.21 são inseridos em íntrons. Assim, o mRNA produzido por esse gene não incluirá nenhuma seqüência de elementos de transposição, pois eles teriam sido recompostos do pré-mRNA com o íntron vizinho. Supostamente, os elementos de transposição inserem-se tanto em éxons quanto em íntrons, mas apenas as inserções em íntrons permanecerão na população porque menos provavelmente causam uma mutação deletéria. As inserções em éxons são ditas sujeitas a **seleção negativa**. Segundo, os humanos, bem como todos os outros organismos multicelulares, podem sobreviver com tanto DNA móvel no genoma porque a maioria é inativa e não pode mover-se nem aumentar o número de cópias. A maioria das seqüências de elementos de transposição em um genoma são remanescentes que se acumularam inativando mutações ao longo do

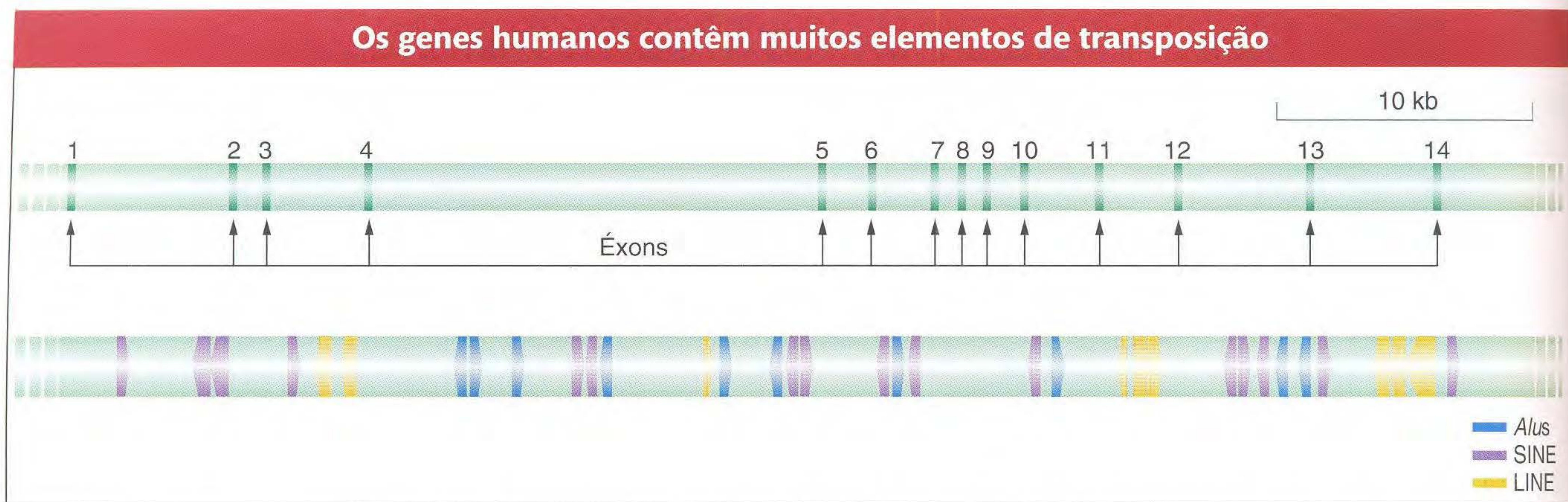


FIG. 14.21 Vários elementos repetitivos são encontrados no gene humano (*HGO*) codificando homogentisato 1,2-dioxigenase, a enzima cuja deficiência causa alcaptonúria. A fileira de cima diagrama as posições dos éxons *HGO*. As localizações de *Alus* (azul), outros SINE (púrpura), e LINE (amarelo) na sequência *HGO* são indicadas na fileira inferior. [De B. Granadino, D. Beltrán-Valero de Bernabé, J.M. Fernández-Cañón, M.A. Peñalva, and S. Rodríguez de Córdoba, "The Human Homogentisate 1,2-dioxigenase (*HGO*) Gene," *Genomics* 43, 1997, 115.]

tempo evolutivo. Outros ainda são capazes de movimento, mas são inativados por mecanismos reguladores do hospedeiro. Existem, entretanto, alguns LINE ativos e *Alus* que conseguiram escapar do controle do hospedeiro e inseriram-se em genes importantes, causando várias doenças humanas. Três inserções separadas de LINE perturbam o gene de fator VIII, causando hemofilia A. Pelo menos 11 inserções *Alu* em genomas humanos foram mostradas causando várias doenças, inclusive hemofilia B (no gene de fator IX), neurofibromatose (no gene *NF1*) e câncer de mama (no gene *BRCA2*).

A frequência geral de mutação espontânea devida à inserção de elementos classe 2 em humanos é muito baixa, correspondendo a menos de 0,2% (1 em 500) de todas as mutações espontâneas caracterizadas. Surpreendentemente, as inserções de retrotransposon correspondem a 10% das mutações espontâneas em outro mamífero, o camundongo. O aumento de aproximadamente 50 vezes nesse tipo de mutação no camundongo mais provavelmente corresponde a uma atividade muito maior desses elementos no genoma de camundongo do que no genoma humano.

Os elementos de transposição compõem a maior fração do genoma humano, com LINE e SINE sendo os mais abundantes. A grande maioria dos elementos de transposição são relíquias que não podem mais mover-se nem aumentar seu número de cópias. Alguns elementos permanecem ativos e seu movimento para os genes pode causar doença.

As gramíneas: retrotransposons LTR desenvolvem-se em grandes genomas

Como já foi mencionado, o paradoxo do valor C é a falta de correlação entre o tamanho do genoma e a complexidade biológica. Como os organismos podem ter um conteúdo gênico muito similar mas diferirem marcadamente no tamanho de seus genomas? Essa situação foi investigada nos cereais. As diferenças nos tamanhos dos genomas dessas gramíneas foram mostradas correlacionando-se primariamente com o número de uma classe

de elementos, os retrotransposons LTR. Os cereais são parentes evolutivos que surgiram de um ancestral comum nos últimos 70 milhões de anos. Assim, seus genomas ainda são muito similares com relação ao conteúdo gênico e organização (chamado de **simetria**; veja os Capítulos 13 e 19), e regiões podem ser diretamente comparadas. Essas comparações revelam que genes ligados no pequeno genoma de arroz estão fisicamente mais próximos do que estão os mesmos genes nos genomas maiores de milho e cevada. Nos genomas de milho e cevada, os genes são separados por grandes grupos de retrotransposons (Figura 14.22).

Abrigos seguros

A abundância de elementos de transposição nos genomas de organismos multicelulares levou alguns pesquisadores a postular que elementos de transposição bem-sucedidos (os que são capazes de atingir um número muito grande de cópias) desenvolveram mecanismos para evitar prejudicar seus hospedeiros não se inserindo em genes hospedeiros. Em vez disso, elementos de transposição bem-sucedidos inserem-se nos chamados **abrigo seguros (safe havens)** no genoma. Para as gramíneas, um abrigo de segurança para novas inserções parece ser em outros retrotransposons. Outro abrigo seguro está na heterocromatina de centrômeros, onde existem poucos genes, mas muito DNA repetitivo (veja o Capítulo 11 para mais informações sobre heterocromatina). Muitas classes de elementos de transposição, tanto em espécies de plantas quanto de animais, tendem a se inserir na heterocromatina centromérica.

Abrigos seguros em genomas pequenos: inserções dirigidas

Em contraste com os genomas de eucariontes multicelulares, o genoma de leveduras unicelulares é muito compacto, com genes próximos e muito poucos íntrons. Com quase 70% de seu genoma como éxons, há uma alta probabilidade de que novas inserções de elementos de transposição perturbem uma sequência codificante. Como já vimos neste capítulo, o genoma de leveduras suporta uma coleção de retrotransposons LTR chamados de elementos *Ty*.

Elementos de transposição em gramíneas são responsáveis por diferenças no tamanho do genoma

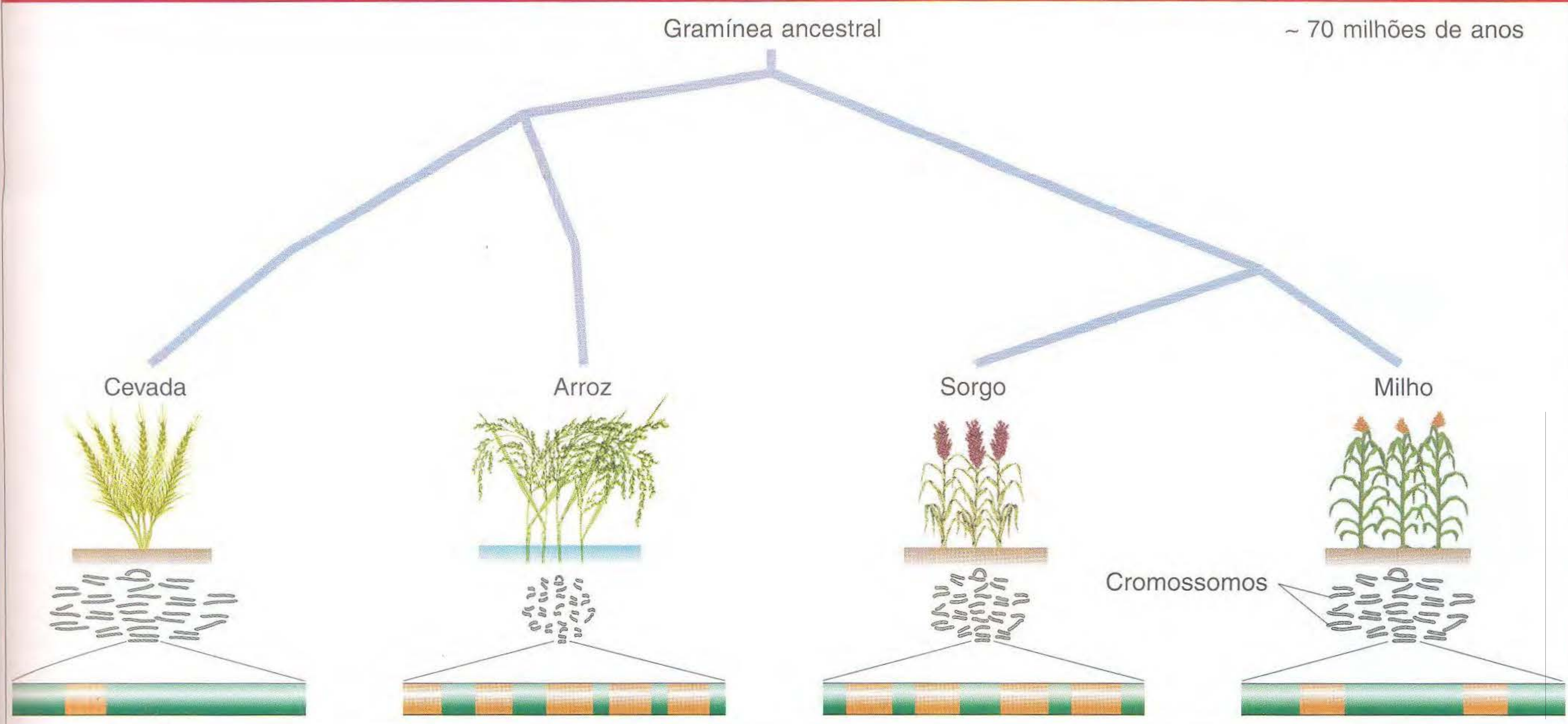


FIG. 14.22 As gramíneas, incluindo cevada, arroz, sorgo e milho, divergiram de um ancestral comum há cerca de 70 milhões de anos. Desde essa época, os elementos de transposição acumularam níveis diferentes de cada espécie. Os cromossomos são maiores em milho e cevada, cujos genomas contêm grandes quantidades de retrotransposons LTR. Verde no genoma parcial na parte de baixo representa um aglomerado de transposons, enquanto laranja representa genes.

Como esses elementos de transposição são capazes de se espalhar para novos sítios nos genomas com poucos abrigos seguros? Os pesquisadores identificaram centenas de elementos *Ty* no genoma sequenciado de leveduras e determinaram que eles não estão distribuídos aleatoriamente, mas sim que cada família de elementos *Ty* insere-se em uma região genômica particular. Por exemplo, a família *Ty3* insere-se quase exclusivamente perto dos genes de tRNA, mas não nesses genes, localizando-se em sítios onde não interferem na produção de tRNA e, supostamente, não prejudicam seus hospedeiros. Os elementos *Ty* desenvolveram um mecanismo que lhes permite inserir-se em regiões particulares do genoma: as proteínas *Ty* necessárias para integração interagem com proteínas específicas de levedura ligadas ao DNA genômico. As proteínas *Ty3*, por exemplo, reconhecem e se ligam a subunidades do complexo de RNA polimerase que foram reunidas em promotores de tRNA (Figura 14.23a).

A habilidade de alguns transposons em se inserir preferencialmente em algumas sequências ou regiões genômicas é chamada de **direcionamento (targeting)**. Um exemplo marcante de direcionamento é ilustrado pelos elementos *R1* e *R2* de artrópodes, inclusive *Drosophila*. *R1* e *R2* são LINE (veja Figura 14.20) que se inserem em genes que produzem RNA ribossômico. Nos artrópodes, várias centenas de genes de rRNA são organizadas em disposições em tandem (Figura 14.23b). Com tantos genes codificando os mesmos produtos, o hospedeiro tolera a inserção em um subgrupo. Entretanto, muitas inserções de *R1* e *R2* mostraram diminuir a viabilidade do inseto, supostamente interferindo na montagem do ribossomo.

Voltando à terapia gênica Este capítulo começou com uma descrição de um distúrbio genético recessivo chamado SCID

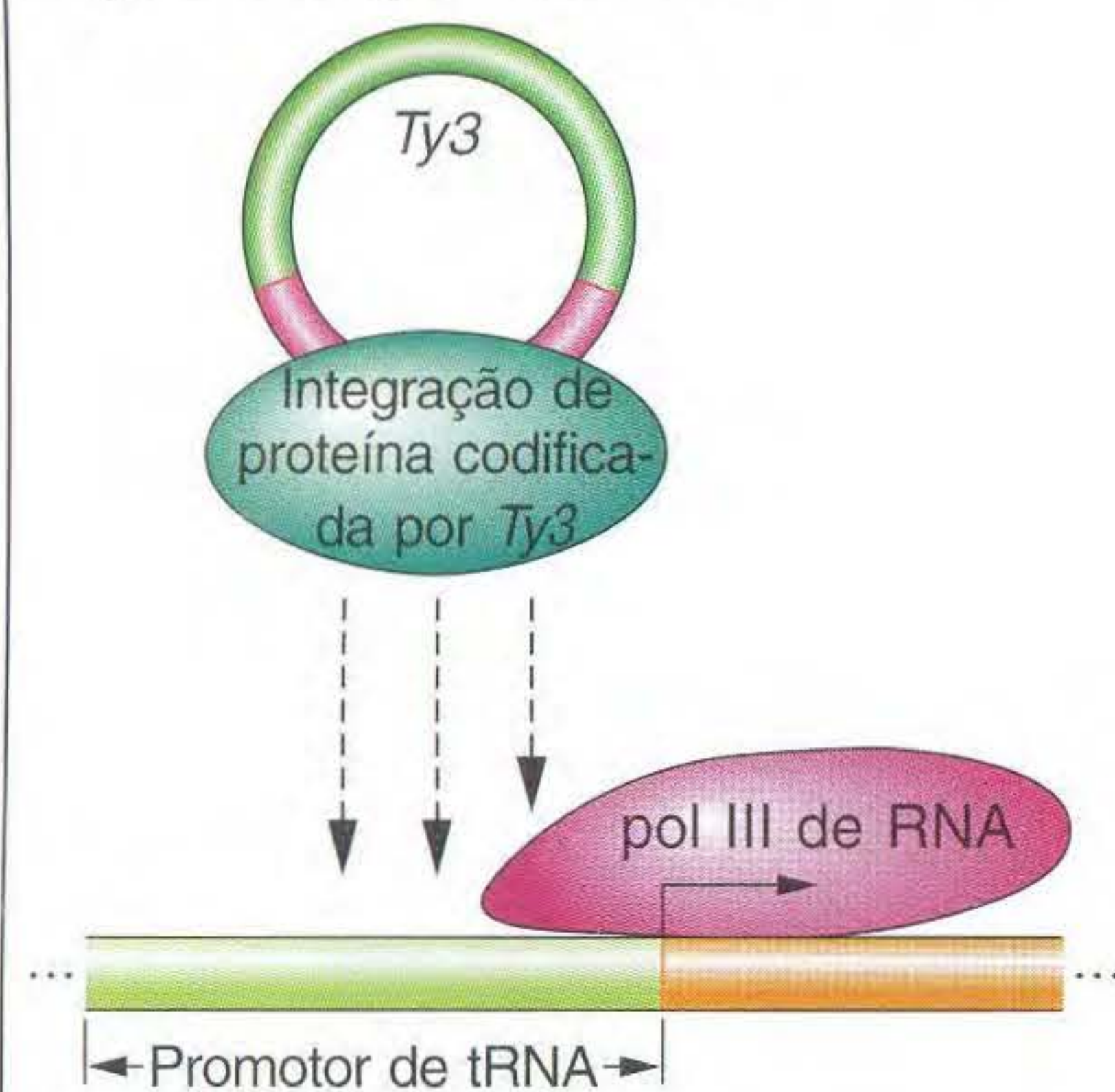
(doença de imunodeficiência combinada grave). Os sistemas imunológicos de pessoas com SCID estão gravemente comprometidos devido a uma mutação em um gene que codifica a enzima adenosina desaminase. Para corrigir esse defeito genético, as células da medula óssea dos pacientes SCID foram colhidas e tratadas com um vetor retroviral contendo um bom gene ADA. As células transformadas foram introduzidas de volta nos pacientes. Os sistemas imunológicos da maioria dos pacientes mostrou uma significativa melhora. Entretanto, a terapia teve um grave efeito colateral: dois dos pacientes desenvolveram leucemia. Em ambos os pacientes, o vetor retroviral inseriu-se (integrou-se) perto de um gene celular cuja expressão aberrante está associada à leucemia. Um cenário provável é que a inserção do vetor retroviral próximo ao gene celular alterou sua expressão e, direta ou indiretamente, causou a leucemia.

Claramente, essa forma de terapia gênica poderia ser muito melhorada se os médicos fossem capazes de controlar onde o vetor retroviral se integra no genoma humano. Já vimos que existem muitas semelhanças entre os retrotransposons LTR e os retrovírus. Esperamos que, entendendo o direcionamento de *Ty* em leveduras, possamos aprender como construir vetores retrovirais que se inserem, e sua carga transgênica, em abrigos de segurança no genoma humano.

Um elemento de transposição bem-sucedido aumenta o número de cópias sem prejudicar o hospedeiro. Um modo pelo qual um elemento aumenta com segurança o número de cópias é direcionar novas inserções para abrigos seguros, regiões do genoma onde existem poucos genes.

Os elementos de transposição inserem-se em abrigos seguros

(a) Integração de *Ty3* em abrigos seguros de gene de tRNA



(b) *R1* e *R2* em abrigos seguros de rRNA

Disposição em tandem de genes de rRNA em cromossomos de insetos

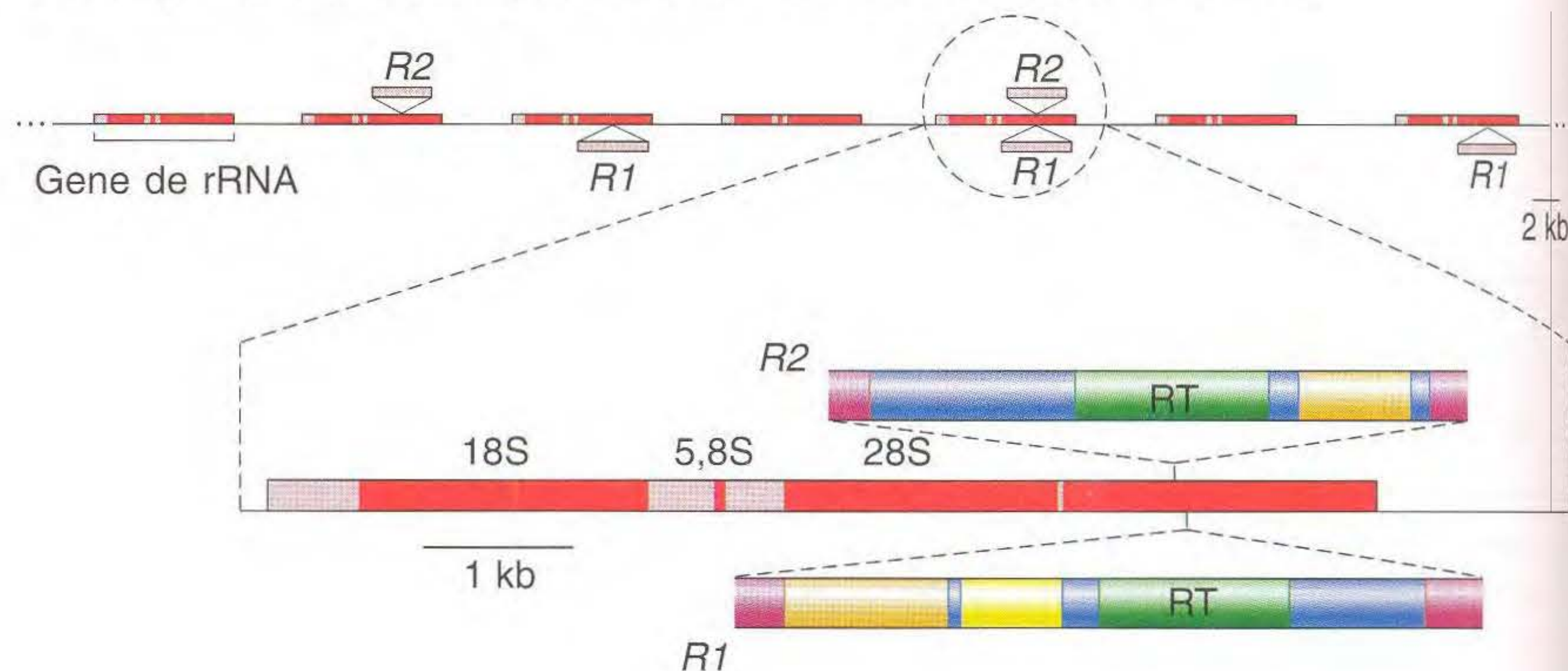


FIG. 14.23 Alguns elementos de transposição são direcionados para abrigos seguros específicos. (a) O retrotransposon da levedura *Ty3* insere-se na região do promotor de genes de RNA transportador. (b) Os retrotransposons não-LTR *R1* e *R2* de *Drosophila* (LINE) inserem-se em genes que codificam RNA ribossômico que são encontrados em longos arranjos em tandem no cromossomo. Apenas os genes de transcriptase reversa (RT) de *R1* e *R2* são representados. [(a) Inspirada em D.F. Voytas and J.D. Boeke, "Ty1 and Ty5 of *Saccharomyces cerevisiae*," in N.L. Craig et al., Eds., *Mobile DNA II*, Chap. 26, Fig. 15, p. 652. ASM Press, 2002. (b) De T.H. Eickbush, "R2 and Related Site-Specific Non-Long Terminal Inverted Repeat Retrotransposons," in N.L. Craig et al., Eds., *Mobile DNA II*, Chap. 34, Fig. 1, p. 814, ASM Press, 2002.]

Resumo

Os elementos de transposição foram descobertos em milho por Barbara McClintock como a causa de várias mutações instáveis. Um exemplo de um elemento não-autônomo é *Ds*, cuja transposição requer a presença do elemento autônomo *Ac* no genoma.

As seqüências de inserção de elementos bacterianos foram os primeiros elementos de transposição isolados molecularmente. Existem muitos tipos diferentes de elementos IS nas linhagens de *E. coli*, e eles estão geralmente presentes em pelo menos várias cópias. Os transposons compostos contêm elementos IS flanqueando um ou mais genes, tais como genes que conferem resistência a antibióticos. Os transposons com genes de resistência podem inserir-se em plasmídeos, sendo então transferidos por conjugação para bactérias não-resistentes.

Existem dois grupos principais de elementos de transposição nos eucariontes: elementos classe 1 (retrotransposons) e elementos classe 2 (transposons de DNA). O elemento *P* foi o primeiro transposon de DNA classe 2 a ser isolado molecularmente. Ele foi isolado de mutações instáveis em *Drosophila* que foram induzidas por disgenesia híbrida. Os elementos *P*

foram desenvolvidos em vetores para a introdução de DNA exógeno nas células germinativas de *Drosophila*.

Ac, *Ds* e *P* são exemplos de transposons de DNA, assim chamados porque o intermediário de transposição é o próprio DNA. Os elementos autônomos tais como *Ac* codificam uma transposase que se liga às pontas de elementos autônomos e não-autônomos e catalisa a excisão do elemento de um sítio doador e a reinserção em um novo sítio-alvo em outra parte do genoma.

Os retrotransposons foram primeiro isolados molecularmente de mutantes de leveduras, e sua semelhança com retrovírus foi imediatamente aparente. Os retrotransposons são elementos classe 1, como são todos os elementos de transposição que usam RNA como seu intermediário de transposição.

Os elementos de transposição ativos isolados de tais organismos-modelo como leveduras, *Drosophila*, *E. coli*, e milho constituem uma fração muito pequena de todos os elementos de transposição no genoma. O seqüenciamento do DNA de genomas inteiros, incluindo o genoma humano, levou ao marcante achado de que quase metade de todo o genoma humano é derivado de elementos de transposição.

Termos-chave

abrigo seguro

Alu

Ativador (*Ac*)

citotipo M

citotipo P

cointegrado

direcionamento

disgenesia híbrida

Dissociação (*Ds*)

duplicação de sítio-alvo

elemento autônomo

elemento classe 1 (retrotransposon)

elemento classe 2 (transposon de DNA)

elemento curto intercalado (SINE)

elemento de seqüência de inserção (IS)

elemento longo intercalado (LINE)

elemento não-autônomo

elemento *P*

elemento tipo *copia*

elemento *Ty*

excisar

fator R

fenótipo instável

longa repetição terminal (LTR)

marcação de transposon

paradoxo do valor *C*

plasmídeo R	seleção negativa	transposição conservativa
provírus	sintenia	transposição replicativa
“recortar e colar”	terapia gênica	transposon (Tn)
repetição invertida (IR)	transcriptase reversa	transposon composto
retrotransposon	transpor	transposon de DNA
retrotransposon de LTR	transposase	transposon simples
retrovírus	transposição	valor C

Problemas Resolvidos

Problema resolvido 1. Os elementos de transposição foram chamados de “genes saltadores” porque parecem saltar de uma posição para outra, deixando o locus antigo e aparecendo em um novo locus. À luz do que hoje sabemos quanto ao mecanismo de transposição, o termo “genes saltadores” é apropriado para os elementos de transposição bacterianos?

SOLUÇÃO

Nas bactérias, a transposição ocorre de dois modos diferentes. O modo conservativo resulta em verdadeiros genes saltadores, porque, nesse caso, o elemento de transposição sai de sua posição original e insere-se em uma nova posição. O outro modo é o modo replicativo. Nessa via, um elemento de transposição move-se para um novo local, replicando-se no DNA-alvo, deixando uma cópia do elemento de transposição no local original. Quando operando pelo modo replicativo, os elementos de transposição não são de fato genes saltadores, pois permanece uma cópia no sítio original.

Problemas

PROBLEMAS BÁSICOS

1. Descreva a produção de plasmídeos de resistência a várias drogas.
2. Descreva resumidamente o experimento que demonstra que a transposição do elemento *Ty1* em leveduras ocorre por um intermediário de RNA.
3. Explique como as propriedades dos elementos *P* em *Drosophila* possibilitam os experimentos de transferência gênica nesse organismo.
4. Embora os elementos classe 2 sejam abundantes nos genomas de eucariontes multicelulares, os elementos classe 1 geralmente constituem a maior fração dos genomas muito grandes, tais como os dos humanos (~ 2.500 Mb), milho (~ 2.500 Mb) e cevada (~ 5.000 Mb). Considerando o que você sabe sobre os elementos classes 1 e 2, o que há nesses mecanismos distintos de transposição que explicaria essa consistente diferença em abundância?
5. Como você viu na Figura 14.21, os genes de eucariontes multicelulares em geral contêm muitos elementos de transposição. Por que a maioria desses elementos não afeta a expressão do gene?
6. O que são abrigos seguros? Existem lugares nos genomas bacterianos mais compactos que podem ser um abrigo seguro para os elementos de inserção?
7. Os Prêmios Nobel geralmente são dados muitos anos após a descoberta. Por exemplo, James Watson, Francis Crick e Maurice Wilkens receberam o Prêmio Nobel de Medicina ou Fisiologia em 1962, quase uma década após sua descoberta da estrutura da dupla hélice do DNA. Entretanto, Barbara McClintock foi premiada em 1983, quase quatro décadas após sua descoberta dos elementos de transposição em milho. Por que você acha que há tanta demora?

PROBLEMAS DESAFIADORES

8. A inserção dos elementos de transposição em genes pode alterar o padrão normal de expressão. Nas situações seguintes, descreva as possíveis consequências na expressão gênica.
 - a. Uma inserção LINE em um acentuador de um gene humano.
 - b. Um elemento de transposição contém um sítio de ligação para um repressor de transcrição e insere-se adjacente a um promotor.
 - c. Um elemento *Alu* insere-se no sítio de corte 3' (AG) de um íntron.
 - d. Um elemento *Ds* que foi inserido no éxon de um gene sai de modo imperfeito e deixa 3 pb atrás no éxon.
 - e. Outra excisão pelo mesmo elemento *Ds* deixa 2 pares de bases atrás no éxon.
 - f. Um elemento *Ds* que foi inserido no meio de um íntron sai de modo imperfeito e deixa 5 pares de bases atrás no íntron.
9. Antes da integração de um transposon, sua transposase faz um corte desencontrado no DNA-alvo do hospedeiro. Se os cortes desencontrados são nos sítios das setas a seguir, desenhe como será a sequência do DNA hospedeiro após o transposon ter sido inserido. Represente o transposon como um retângulo.

↓
AATTTGG CCTAGTACTAATTGGTTGG
TTAAACCGGATCATGATTAACCAACC
↑

10. Em *Drosophila*, M. Green descobriu um alelo *singed* (*sn*) com algumas características incomuns. As fêmeas homozigotas para esse alelo ligado ao X têm cerdas *singed*, mas várias áreas de cerdas *sn*⁺ (tipo selvagem) em sua cabeça,

tórax e abdome. Quando essas moscas são cruzadas com machos *sn*, algumas fêmeas dão apenas prole *singed*, mas outras dão tanto *singed* quanto prole tipo selvagem em proporções variáveis. Explique esses resultados.

11. Considere duas plantas de milho:

a. Genótipo C/c^m ; Ac/Ac^+ , onde c^m é um alelo instável causado pela inserção de *Ds*.

b. Genótipo C/c^m , onde c^m é um alelo instável causado pela inserção de *Ac*.

Que fenótipos seriam produzidos e em que proporções quando (1) cada planta é cruzada com um mutante de substituição de par c/c e (2) a planta na parte *a* é cruzada com a planta na parte *b*? Suponha que *Ac* e *c* não estão ligados, que a frequência de quebras cromossômicas é desprezível e que o mutante c/C é Ac^+ .

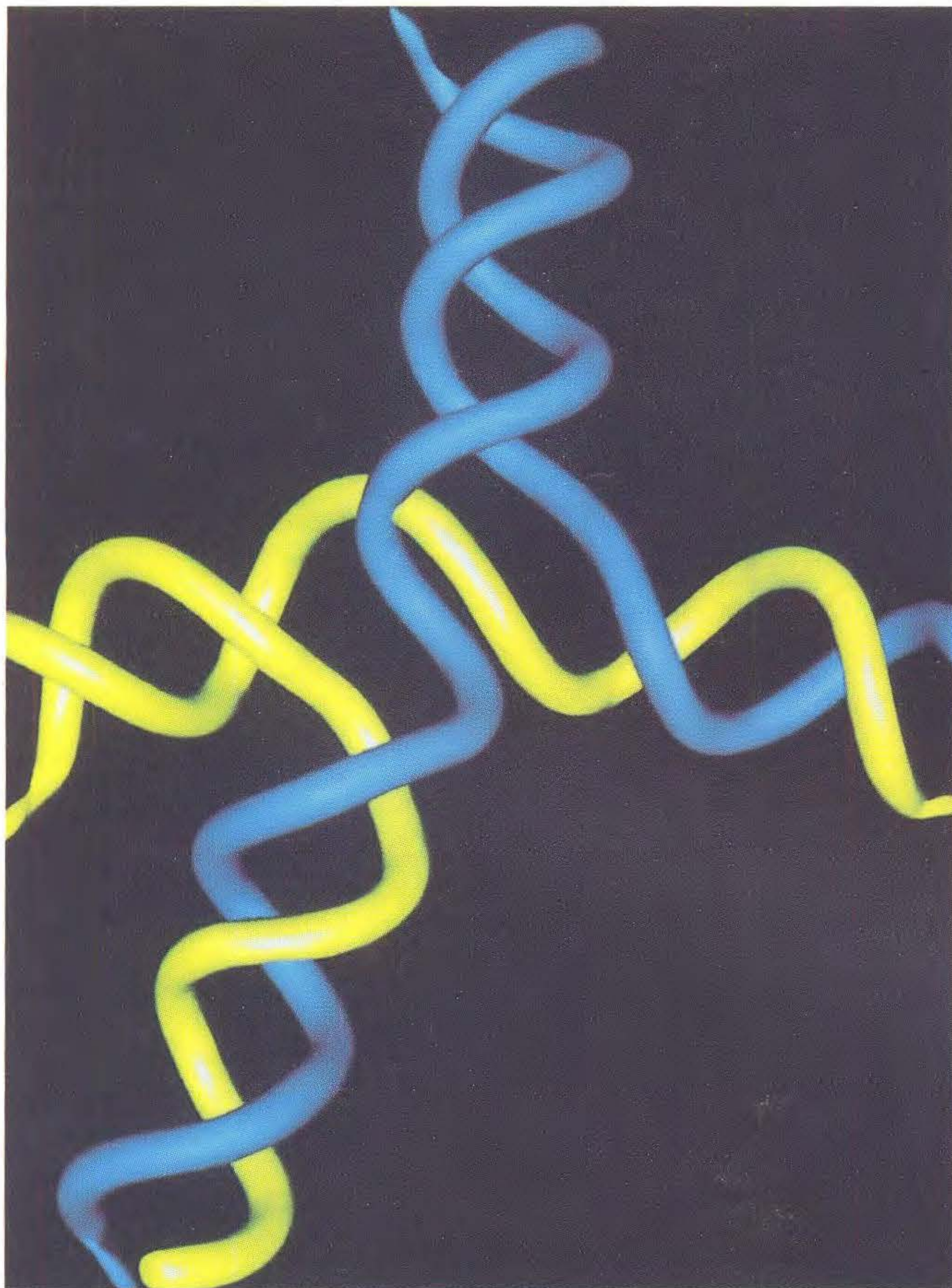
12. Você encontra sua amiga, uma cientista, em um ginásio e ela começa a lhe falar sobre um gene de camundongo que ela está estudando no laboratório. O produto desse

gene é uma enzima necessária para tornar a pelagem marrom. O gene é chamado de *FB* e a enzima é chamada de FB. Quando *FB* é mutante e não pode produzir a enzima FB, a pelagem é branca. A cientista lhe diz que ela isolou o gene de dois camundongos com pelagem marrom e que, surpreendentemente, descobriu que os dois genes diferem pela presença de um SINE de 250 pb (como o elemento humano *Alu*) no gene *FB* de um camundongo, mas não no gene do outro. Ela não compreende como é possível essa diferença, sobretudo porque determinou que ambos os camundongos fazem a enzima FB. Você pode ajudá-la a formular uma hipótese que explique por que o camundongo ainda pode produzir a enzima FB com um elemento de transposição em seu gene *FB*?

13. O genoma de levedura tem elementos classe 1 (*Ty1*, *Ty2*, e assim por diante), mas não elementos classe 2. Você pode pensar em um motivo possível para que os elementos de DNA não tenham tido sucesso no genoma de levedura?

Mutação, Reparo e Recombinação

15



Um modelo de computação de uma junção de Holliday. [Julie Newdol, Computer Graphics Laboratory, University of California, San Francisco. Copyright by Regents, University of California.]

Uma paciente jovem desenvolve muitos pequenos crescimentos, tipo sarda, pré-cancerosos na pele, e é extremamente sensível à luz do sol (Figura 15.1). Uma história familiar é colhida, e a paciente é diagnosticada com uma doença recessiva autossômica chamada xeroderma pigmentoso. Durante sua vida, ela estará propensa a desenvolver cânceres pigmentados de pele. Como podemos compreender os muitos efeitos exercidos pela mutação do xeroderma pigmentoso? Vários genes diferentes podem ser mutados para gerar o fenótipo do xeroderma pigmentoso. Em uma pessoa sem a doença, cada um desses genes contribui para processos bioquímicos na célula que respondem a danos químicos ao DNA e reparo desse dano antes que leve à formação de novas mutações. Mais adiante neste capítulo, veremos como as mutações no sistema de reparo levam a doenças genéticas, tais como o xeroderma pigmentoso.

Perguntas Fundamentais

- Qual a natureza molecular das mutações?
- Qual a base das mutações espontâneas em comparação com as induzidas?
- Quais são os mecanismos de reparo biológico?
- Que doenças genéticas humanas são causadas por mutações nos mecanismos de reparo?
- Como a evolução explorou as vias de reparo em processos celulares normais, tais como as recombinações meióticas?

Tópicos

- 15.1 As consequências fenotípicas das mutações no DNA
- 15.2 A base molecular das mutações espontâneas
- 15.3 A base molecular das mutações induzidas
- 15.4 Mecanismos biológicos de reparo
- 15.5 O mecanismo de crossing over meiótico
- 15.6 Câncer: uma consequência fenotípica importante da mutação

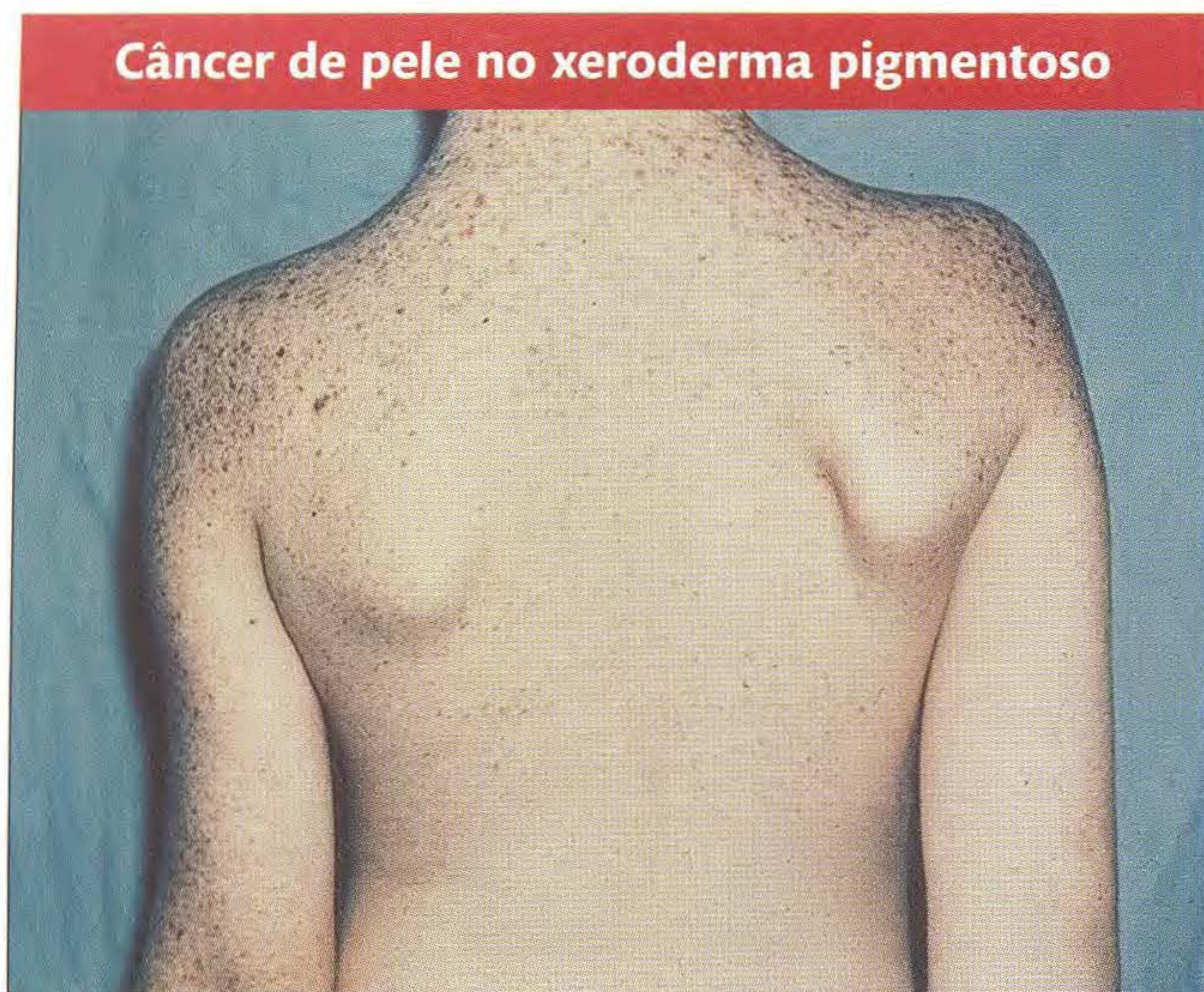


FIG. 15.1 A doença herdada recessivamente xeroderma pigmentoso é causada por deficiência em uma enzima que ajuda a corrigir DNA danificado. Essa deficiência enzimática leva à formação de cânceres de pele com a exposição da pele aos raios ultravioleta da luz do sol. [Ken Greer/Visuals Unlimited.]

As pessoas com xeroderma pigmentoso são exemplos de *variantes* genéticas que apresentam diferenças fenotípicas em uma ou mais características particulares. Como a genética é o estudo das diferenças herdadas, a análise genética não seria possível sem variantes. Os capítulos precedentes incluem muitas análises da herança de tais variantes. Consideraremos agora sua origem. Como surgem as variantes genéticas?

Dois processos importantes são responsáveis pela variação genética: *mutação* e *recombinação*. Vimos que a mutação é uma mudança na sequência de DNA de um gene. A mutação é especialmente significativa porque é fonte primária da mudança evolutiva; novos alelos surgem em todos os organismos, alguns espontaneamente e outros resultantes de exposição a radiação ou substâncias químicas no ambiente. Os novos alelos produzidos por mutação tornam-se a matéria-prima para um segundo nível de variação, efetuado por recombinação. Como seu nome sugere, a recombinação é o resultado de processos celulares que fazem com que alelos de genes diferentes se tornem agrupados em novas combinações (veja o Capítulo 4). Para usar uma analogia, a mutação produz novas cartas de baralho, e, então, a recombinação as mistura e cria novos conjuntos.

No ambiente celular, as moléculas de DNA não são absolutamente estáveis; cada par de bases em uma dupla hélice de DNA tem certa probabilidade de mutar. Como veremos, o termo *mutação* cobre uma ampla gama de tipos diferentes de mudanças. Essas mudanças variam de uma simples troca de um par de bases por outro até o desaparecimento de um cromossomo inteiro. No Capítulo 16, consideraremos as alterações mutacionais que afetam cromossomos inteiros ou grandes pedaços de cromossomos. No presente capítulo, enfocaremos eventos mutacionais que ocorrem *dentro* de genes individuais. Chamaremos esses eventos de *mutações gênicas*.

As células desenvolveram sistemas sofisticados para identificar e reparar DNA danificado, impedindo assim a ocorrência de mutações. Podemos ver o DNA como sendo submetido a um dinâmico cabo-de-guerra entre os processos que danificam o DNA e levam a novas mutações e os processos químicos de reparo celular que constantemente monitoram o DNA quanto a tais danos e os corrigem. Entretanto, esse cabo-de-guerra não é direto. Como já mencionado, as mutações fornecem a matéria-prima para a evolução, e, portanto, a introdução de um baixo nível de mutação deve ser tolerado. Veremos que os sistemas de replicação do DNA e reparo podem introduzir mutações. Outros transformam mutações potencialmente devastadoras (tais como quebras bifilamentares) em mutações que podem afetar apenas um único produto gênico.

Veremos que a classe potencialmente mais séria de dano ao DNA, uma quebra bifilamentar, também é uma etapa intermediária no processo celular normal de recombinação pelo crossing over meiótico. Assim, podemos traçar paralelos entre a mutação e a recombinação em dois níveis. Primeiro, como já foi mencionado, a mutação e a recombinação são as principais fontes de variação. Segundo, os mecanismos de reparo do DNA e recombinação têm algumas características em comum, incluindo o uso de algumas das mesmas proteínas. Por esse motivo, exploraremos os mecanismos de reparo do DNA primeiro e, então, os compararemos ao mecanismo de recombinação de DNA.

► 15.1 As Consequências Fenotípicas das Mutações no DNA

O termo **mutação de ponto** refere-se tipicamente à alteração de um único par de bases do DNA ou um pequeno número de pares de bases adjacentes. Nesta seção, consideraremos os efeitos de tais mudanças a nível fenotípico. As mutações de ponto são classificadas em termos moleculares na Figura 15.2, que mostra os principais tipos de mudanças de DNA e seus efeitos na função da proteína onde elas ocorrem dentro da região codificante de um gene.

Tipos de mutação de ponto

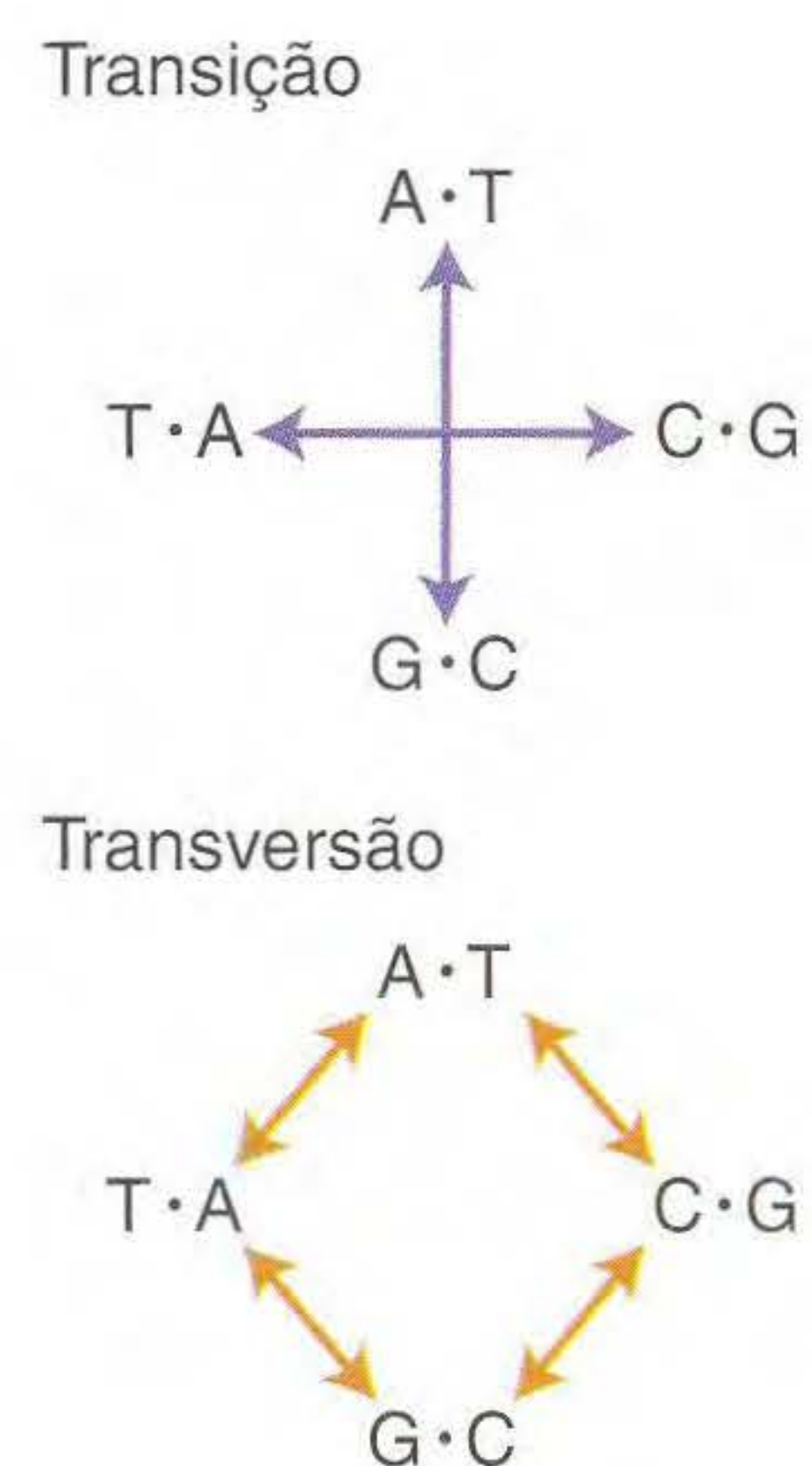
Os dois principais tipos de mutação de ponto no DNA são as *substituições de bases* e as *inserções* ou *deleções*. As substituições de bases são mutações nas quais um par de bases é substituído por outro. As substituições de bases podem ser divididas em dois subtipos: *transições* e *transversões*. Para descrevermos esses subtipos, consideraremos como uma mutação altera a sequência de um filamento de DNA (a mudança complementar ocorrerá no outro filamento). Uma **transição** é a substituição de uma base por outra base da mesma categoria química. Tanto uma purina é substituída por outra purina (de A para G ou de G para A) quanto uma pirimidina é substituída por outra pirimidina (de C por T ou de T por C). Uma **transversão** é o oposto, a substituição de uma base de categoria química por uma base de outra. Tanto

Consequências de mutações de ponto dentro de genes

Tipos de mutações no DNA	Resultados em nível molecular			
Sem mutação		<div>TreLisArgGli</div> <div>Códon 1Códon 2Códon 3Códon 4</div> <div>Tipo selvagem</div> <div>A C A A A G A G A G G T</div> <div>Códons que especificam proteína tipo selvagem.</div>		
Transição ou transversão	Mutação sinônima	<div>TreLisArgGli</div> <div>A C A A A G A G C G G T</div> <div>Códon alterado especifica o mesmo aminoácido.</div>		
	Mutação de sentido trocado (conservativa)	<div>TreLisLisGli</div> <div>A C A A A G A A A G G T</div> <div>Códon alterado especifica uma substância similar a um aminoácido.</div>		
	Mutação de sentido trocado (não-conservativa)	<div>TreLisIleGli</div> <div>A C A A A G A T A G G T</div> <div>Códon alterado especifica um aminoácido quimicamente diferente.</div>		
	Mutação sem sentido	<div>TreFIM</div> <div>A C A T A G A G A G G T</div> <div>Códon alterado indica término da cadeia.</div>		
Indel				
Inserção de bases	Mudança de matriz de leitura	<div>TreGluGluArg...</div> <div>A C A G A A G A G A G G T ...</div>		
Deleção de base	Mudança de matriz de leitura	<div>TreArgGluVal...</div> <div>A C A A G A G A G G T ...</div> <div>A</div>		

FIG. 15.2 Uma mutação de ponto dentro da região codificante de um gene varia em seus efeitos no funcionamento da proteína. As proteínas com mutações sinônimas e de sentido trocado geralmente ainda são funcionais.

uma pirimidina é substituída por uma purina (de C por A, C por G, T por A ou T por G) quanto uma purina é substituída por uma pirimidina (de A por C, A por T, G por C ou G por T). Ao descrevermos as mesmas mudanças ao nível bifilamentar do DNA, devemos representar ambos os membros de um par de bases no mesmo local relativo. Assim, um exemplo de uma transição é G · C → A · T; de uma transversão é G · C → T · A. As mutações de inserção ou deleção são na verdade inserções ou deleções de pares de *nucleotídeos*. Entretanto, a convenção é chamá-las inserções ou deleções de pares de *bases*. Coletivamente, elas são chamadas de *mutações indel* (para *inserção-deleção*). A mais simples dessas mutações é a adição ou deleção de um único par de bases. As mutações às vezes surgem pela adição ou deleção simultânea de vários pares de bases de uma vez. Como veremos mais adiante neste capítulo, os mecanismos que seletivamente produzem alguns tipos de adições ou deleções de vários pares de bases são a causa de algumas doenças genéticas humanas.



As consequências moleculares das mutações de ponto em uma região codificante

Quais são as consequências funcionais desses tipos diferentes de mutações de ponto? Primeiro, considere o que ocorre quando uma mutação surge em uma parte codificante de polipeptídeos de um gene. Para substituições de uma única base, existem vários resultados possíveis, mas todos são consequências diretas de dois aspectos do código genético: redundância do código e existência de códons de término de tradução (veja Figura 15.2).

- **Mutações sinônimas.** A mutação muda um códon de um aminoácido por outro códon desse mesmo aminoácido. As mutações sinônimas também são chamadas de mutações *silenciosas*.
- **Mutações de sentido trocado.** O códon para um aminoácido é trocado por um códon para outro aminoácido. As mutações de sentido trocado às vezes são chamadas de mutações *não-sinônimas*.
- **Mutações sem sentido.** O códon para um aminoácido é mudado para um códon de término de tradução (fim).

As substituições sinônimas nunca alteram a sequência de aminoácidos da cadeia polipeptídica. A gravidade do efeito das mutações de sentido trocado e sem sentido no polipeptídeo diferem de caso para caso. Por exemplo, uma mutação de sentido trocado pode substituir um aminoácido por outro aminoácido quimicamente similar, chamada de **substituição conservativa**. Nesse caso, a alteração menos provavelmente afeta gravemente a estrutura e função da proteína. Alternativamente, um aminoácido pode ser substituído por um aminoácido quimicamente diferente em uma **substituição não-conservativa**. Esse tipo de alteração mais provavelmente produz uma grave mudança na estrutura e função da proteína. As mutações sem sentido levarão a um término prematuro da tradução. Assim, elas têm um efeito considerável no funcionamento da proteína. Quanto mais perto a mutação sem sentido estiver da ponta 3' da matriz de leitura aberta (ORF), mais plausível é que a proteína resultante possa ter alguma atividade biológica. Entretanto, muitas mutações sem sentido produzem proteínas completamente inativas.

Mudanças de um único par de bases que inativam proteínas em geral são devidas a mutações de sítio de corte. Como visto na Figura 15.3, tais mudanças podem alterar marcadamente as regiões codificantes, levando a grandes inserções ou deleções que podem ser ou não na matriz de leitura.

Como as mutações sem sentido, as mutações indel podem ter consequências na sequência de polipeptídeos que se estendem bem além do sítio da própria mutação (veja Figura 15.2). Lembre que a sequência de mRNA é "lida" pelo aparelho de tradução na matriz de leitura, três bases de cada vez (um códon). A adição ou deleção de um único par de bases de DNA muda a matriz de leitura do restante do processo de tradução, a partir do sítio da mutação do par de bases até o próximo códon de fim na nova matriz de leitura. Assim, essas lesões são chamadas de **mudanças de matriz de leitura**. Essas mutações fazem com que toda a sequência de aminoácidos seguintes ao sítio mutante não tenha relação com a sequência original de aminoácidos. Assim, a mudança de matriz de leitura resulta tipicamente na perda completa da estrutura e função normal da proteína.

As consequências moleculares das mutações de ponto em uma região não-codificante

Vejamos agora as mutações que ocorrem em sequências reguladoras e outras não-codificantes. Essas partes de um gene que não codificam diretamente uma proteína contêm muitos sítios cruciais de ligação do DNA para proteínas intercaladas entre as sequências que não são essenciais para a expressão ou atividade gênica. Ao nível do DNA, os sítios de ligação incluem os sítios aos quais se ligam a RNA polimerase e seus fatores associados, bem como sítios aos quais se ligam proteínas específicas reguladoras da transcrição. Ao nível do RNA, outros sítios de ligação importantes incluem os sítios de ligação de ribossomos de mRNA bacterianos, os sítios de corte 5' e 3' para união de éxons em mRNA eucarióticos e sítios que regulam a tradução e localizam o mRNA em áreas particulares e compartimentos dentro da célula.

As ramificações de mutações em partes de um gene que não os segmentos codificantes de polipeptídeos são mais difíceis de prever do que as mutações nos segmentos codificantes. Em geral, as consequências funcionais de qualquer mutação de ponto em tal região dependem de a mutação

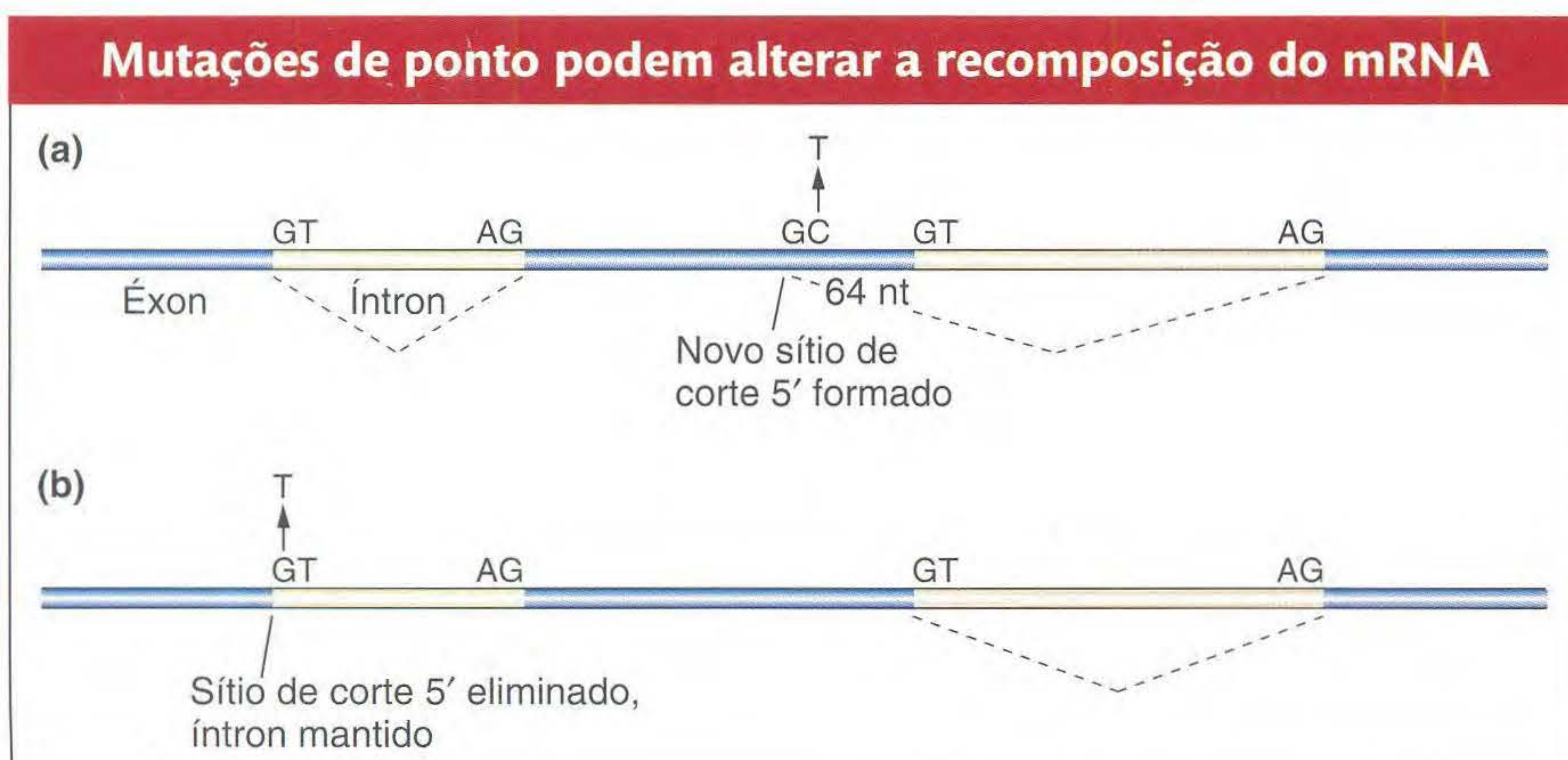


FIG. 15.3 Dois exemplos mostram as consequências de mutações de ponto em sítios de corte. (a) Uma mutação por transição de C para T leva a um dinucleotídeo GT no éxon, formando um novo sítio de corte em 5'. Como resultado, 64 nucleotídeos no final de um éxon são removidos. (b) Uma mutação por transversão de G para T eliminaria o sítio de corte em 5'; logo, o íntron seria retido no mRNA.

Consequências das mutações de ponto nos produtos gênicos

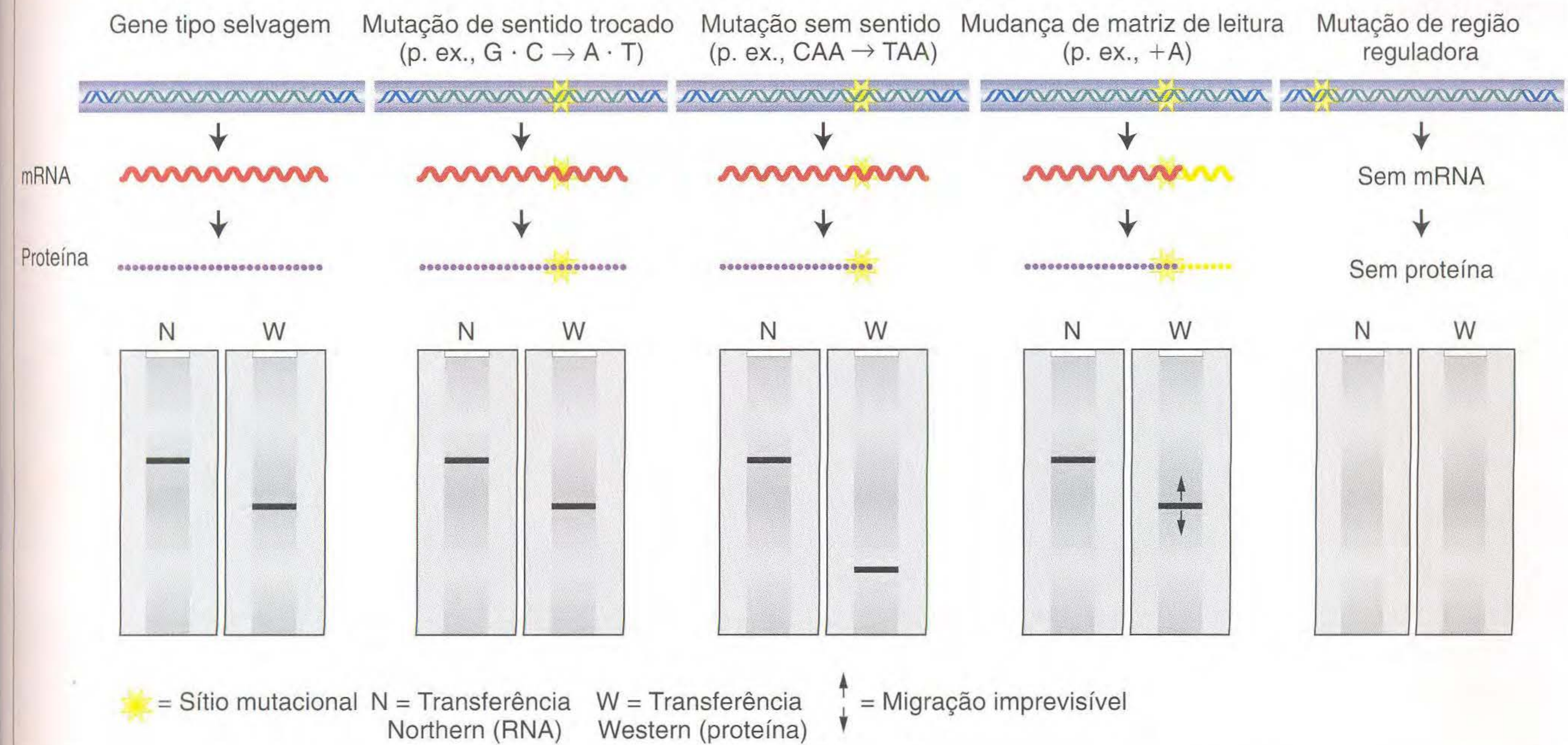


FIG. 15.4 Mutações de ponto em regiões codificantes podem alterar a estrutura da proteína com ou sem alteração no tamanho do mRNA. As mutações de ponto em regiões reguladoras podem evitar a síntese de mRNA (e proteína).

perturbar (ou criar) um sítio de ligação. As mutações que perturbam esses sítios têm o potencial de mudar o padrão de expressão de um gene alterando a quantidade do produto expresso em determinado momento ou em determinado tecido, ou alterando a resposta a alguns estímulos ambientais. Tais mutações reguladoras irão alterar a quantidade do produto protéico, mas *não* a estrutura da proteína. Alternativamente, algumas mutações de sítios de ligação podem bloquear completamente uma etapa necessária na expressão de um gene normal (tal como a ligação da RNA polimerase ou fatores de recomposição) e, assim, inativar totalmente o produto gênico ou bloquear sua formação. A Figura 15.4 mostra alguns exemplos de como tipos diferentes de mutações afetam o mRNA e as proteínas.

É importante ter em mente a distinção entre a ocorrência de uma mutação gênica — isto é, uma mudança na sequência de DNA de determinado gene — e a detecção de tal evento ao nível fenotípico. Muitas mutações de ponto dentro das sequências não-codificantes provocam pouca ou nenhuma mudança fenotípica. Essas mutações estão situadas entre os sítios de ligação de DNA para proteínas regulatórias. Tais sítios podem ser funcionalmente irrelevantes ou outros sítios dentro do gene podem duplicar sua função.

15.2 A Base Molecular das Mutações Espontâneas

As mutações gênicas podem surgir espontaneamente ou ser induzidas. As **mutações espontâneas** são mutações de ocorrência natural e surgem em todas as células. As **mutações induzidas** surgem pela ação de alguns agentes, chamados de

mutágenos, que aumentam a taxa com a qual as mutações ocorrem. Nesta seção, consideraremos a natureza das mutações espontâneas.

Teste de flutuação de Luria e Delbrück

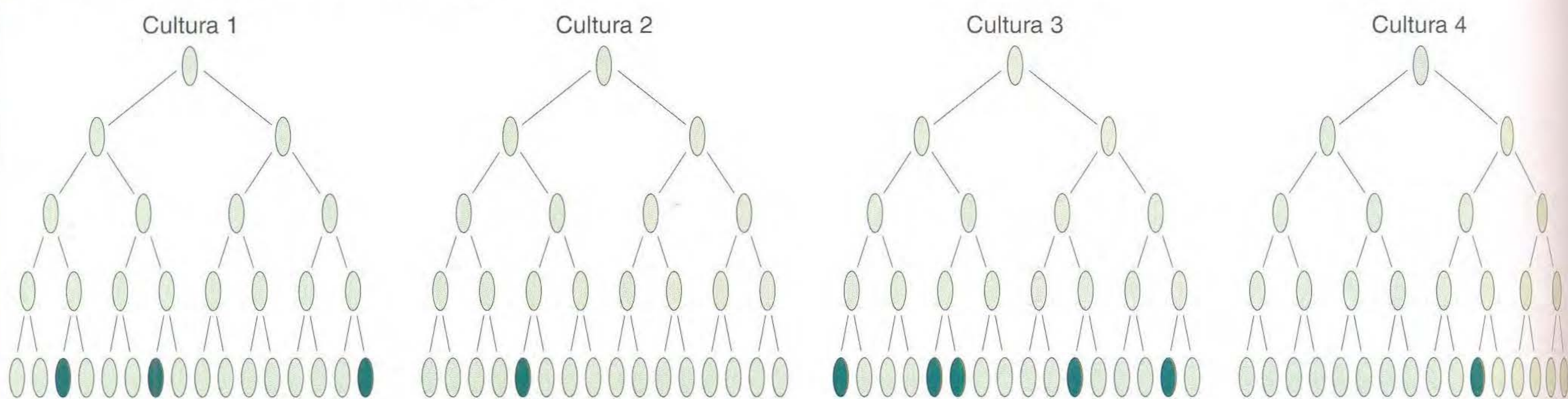
A origem das mudanças hereditárias espontâneas sempre foi um tópico de considerável interesse. Entre as primeiras perguntas feitas pelos geneticistas estava: As mutações espontâneas ocorrem em resposta a estímulos externos ou são variantes presentes em baixa frequência na maioria das populações? Um sistema experimental ideal para abordar essa importante pergunta foi a análise de mutações em bactérias que conferem resistência a agentes ambientais específicos normalmente não tolerados por células tipo selvagem.

Um experimento de Salvador Luria e Max Delbrück, em 1943, foi *particularmente influenciador em modelar nossa compreensão sobre a natureza da mutação*, não só em bactérias, como também em organismos de modo geral. Na época, já se sabia que, se bactérias *E. coli* são dispersas em uma placa de meio nutriente na presença do fago T1, os fagos logo infectam e matam as bactérias. Entretanto, raramente, mas regularmente, foram vistas colônias resistentes ao ataque de fagos. Essas colônias eram estáveis e pareciam ser mutantes genuínos. Entretanto, se esses mutantes eram produzidos espontânea mas aleatoriamente no tempo, ou se a presença do fago induzia uma mudança fisiológica que causava resistência, nós não sabíamos.

Luria raciocinou que, se as mutações ocorriam espontaneamente, então as mutações deveriam ocorrer em tempos diferentes em culturas diferentes. Nesse caso, os números de colônias resistentes por cultura deveriam apresentar alta

Hipótese do “teste de flutuação”

(a) Mudança fisiológica induzida por fago



(b) Mutação aleatória

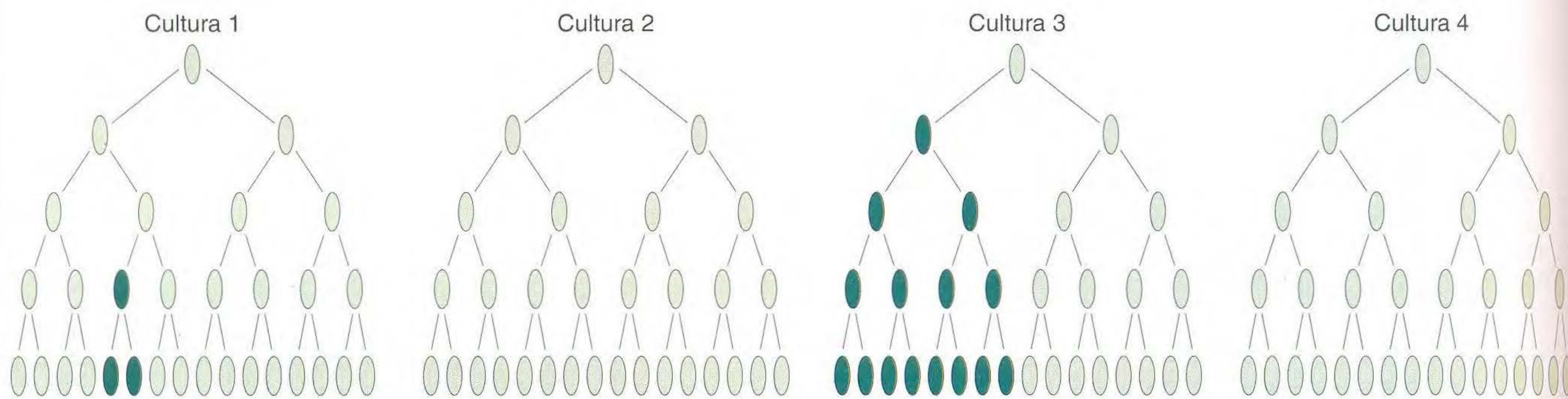


FIG. 15.5 Os pedigrees de células ilustram as expectativas de duas hipóteses contrastantes sobre a origem de células resistentes. [De G.S. Stent and R. Calendar, *Molecular Genetics*, 2nd ed. W.H. Freeman and Company, 1978.]

variação (ou “flutuação”, em suas palavras). Depois, ele disse que a idéia lhe surgiu quando observou os retornos flutuantes obtidos por colegas jogando em uma máquina em um local de dança em um *country club*; donde a origem do termo “*jackpot*” mutation.

Luria e Delbrück projetaram seu “**teste de flutuação**” do seguinte modo. Eles inocularam 20 pequenas culturas, cada uma com poucas células, e as incubaram até ter 10^8 células por mililitro. Ao mesmo tempo, uma cultura muito maior também foi inoculada e incubada até haver 10^8 células por mililitro. As 20 culturas individuais e 20 amostras do mesmo tamanho foram plaqueadas na presença do fago. As 20 culturas individuais mostraram alta variação no número de colônias resistentes: 11 placas tinham 0 colônias resistentes, e o restante tinha 1, 1, 3, 5, 5, 6, 35, 64 e 107 por placa (Figura 15.5a). As 20 amostras da grande cultura mostraram muito menos variação de placa a placa, todas na faixa de variação de 14 a 26. Se o fago estivesse induzindo mutações, não haveria motivo por que a flutuação deveria ser mais alta nas culturas individuais, pois todas foram igualmente expostas ao fago. A melhor explicação é que a mutação estava ocorrendo aleatoriamente no tempo: as primeiras mutações davam números maiores de células resistentes porque as células mutantes tiveram tempo para produzir muitas descendentes resistentes. As últimas mutações produziram

menos células resistentes (Figura 15.5b). Esse resultado levou ao “paradigma” da mutação; isto é, seja em vírus, bactérias ou eucariontes, a mutação pode ocorrer em qualquer célula em qualquer tempo, e sua ocorrência é ao acaso. Por esse e outro trabalho, Luria e Delbrück receberam o Prêmio Nobel em 1969. Curiosamente, isso foi após seu primeiro estudante de graduação, James Watson, ganhar seu Prêmio Nobel (com Francis Crick em 1964) pela descoberta da estrutura da dupla hélice do DNA.

Essa análise elegante sugere que as células resistentes são selecionadas pelo agente ambiental (aqui, o fago), e não produzidas por elas. A existência de mutantes em uma população pode ser demonstrada diretamente antes da seleção? Essa demonstração foi possibilitada pelo uso de uma técnica chamada de **plaqueamento réplica**, desenvolvida por Joshua e Esther Lederberg em 1952. Uma população de bactérias foi plaqueada em meio não-seletivo — isto é, meio não contendo fago —, e cresceu uma colônia de cada célula. Essa placa foi chamada de *placa master*. Um pedaço estéril de veludo foi pressionado levemente sobre a superfície da placa master, e o veludo captou células onde havia colônias (Figura 15.6). Desse modo, o veludo captou uma “impressão” da colônia de toda a placa. O veludo foi então tocado nas placas réplica contendo meio seletivo (isto é, contendo fago T1). Ao tocar as placas de veludo, as células grudadas ao veludo são inoculadas nas placas réplica nas mesmas posições

Plaqueamento réplica

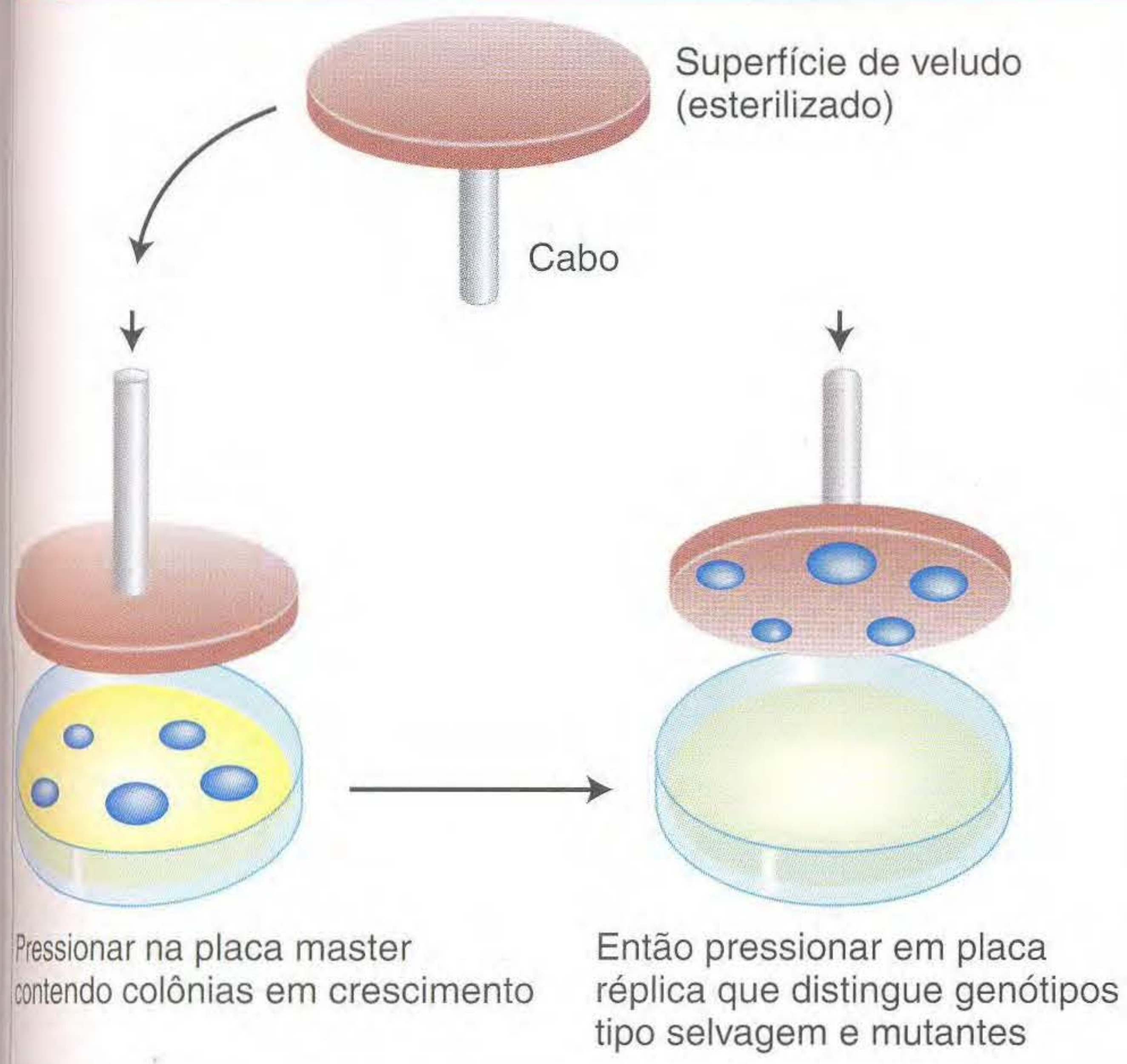
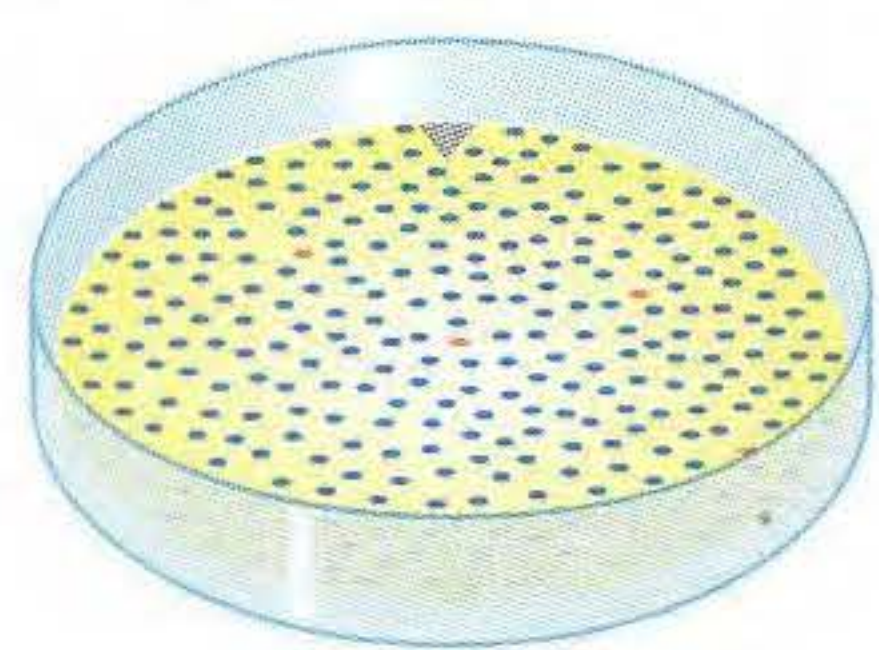


FIG. 15.6 O plaqueamento réplica revela colônias mutantes em uma placa master através de seu comportamento em placas réplica seletivas. [De G.S. Stent and R. Calendar, *Molecular Genetics*, 2nd ed. W.H. Freeman and Company, 1978.]

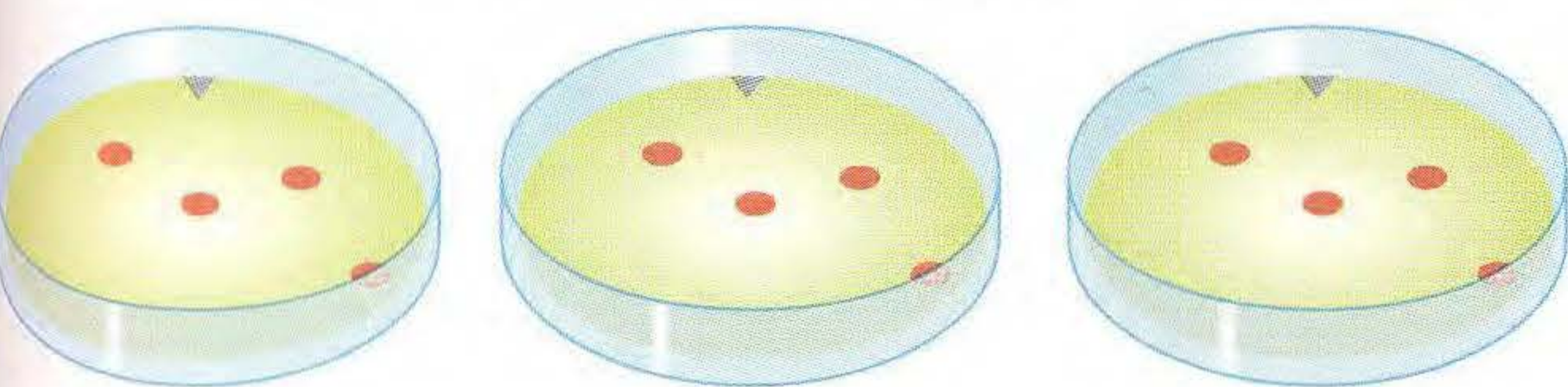
relativas às das colônias da placa master original. Como esperado, raras colônias mutantes resistentes foram encontradas nas placas réplica, mas as múltiplas placas réplica mostraram padrões idênticos de colônias resistentes (Figura 15.7). Se as

O plaqueamento réplica demonstra a presença de mutantes antes da seleção

Placa master contendo 10^7 colônias de *Ton^s E. coli* (sensível a T1)



Plaqueamento réplica



Placa 1 Placa 2 Placa 3

Série de placas réplica contendo altas concentrações de fago T1 e quatro colônias *Ton^r*

FIG. 15.7 Os padrões idênticos nas réplicas mostram que as colônias resistentes são da placa master. [De G.S. Stent and R. Calendar, *Molecular Genetics*, 2nd ed. W.H. Freeman and Company, 1978.]

mutações tivessem ocorrido *após* a exposição aos agentes seletivos, os padrões para cada placa seriam tão aleatórios quanto as próprias mutações. Os eventos de mutação devem ter ocorrido antes da exposição ao agente seletivo. Novamente, esses resultados confirmam que a mutação ocorre aleatoriamente o tempo todo, e não em resposta a um agente seletivo.

A mutação é um processo aleatório. Qualquer alelo em qualquer célula pode mutar a qualquer momento.

Mecanismos de mutações espontâneas

As mutações espontâneas surgem de uma variedade de fontes. Uma fonte é o processo de replicação de DNA. Embora a replicação de DNA seja um processo marcadamente preciso, são cometidos erros na cópia de milhões (e mesmo bilhões) de pares de bases em um genoma. As mutações espontâneas também surgem em parte porque o DNA é uma molécula muito frágil, e o próprio ambiente celular pode danificá-la. Como descrito no Capítulo 14, as mutações podem ser causadas pela inserção de um elemento de transposição de outro local do genoma. Neste capítulo, enfocamos as mutações que não são causadas por elementos de transposição.

Erros na replicação do DNA Um erro na replicação do DNA pode ocorrer quando se forma um par errado de nucleotídeos (digamos, A-C) na síntese de DNA, levando a uma substituição de bases que pode ser uma transição ou uma transversão. Outros erros podem adicionar ou subtrair pares de bases de modo que é criada uma mudança de matriz de leitura.

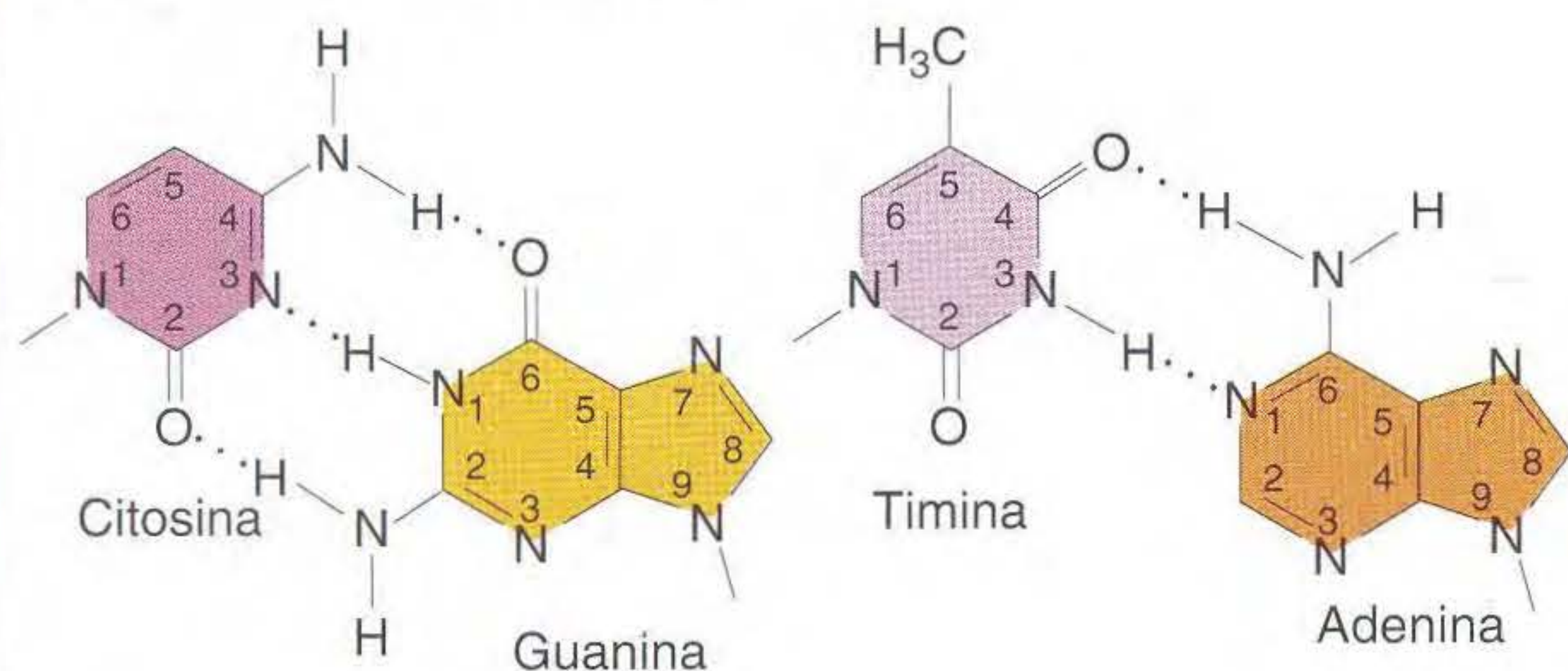
Transições Cada uma das bases do DNA apresenta-se em uma ou várias formas, chamadas de **tautômeros**, que são isômeros que diferem nas posições de seus átomos e nas ligações entre os átomos. As formas estão em equilíbrio. A forma **ceto** de cada base está normalmente presente no DNA (Figura 15.8a), enquanto as formas **imino** e **enol** das bases são raras. As formas imino e ceto podem parear com a base errada, formando um mal pareamento. A habilidade do tautômero errado em fazer um mal pareamento e causar uma mutação no curso da replicação do DNA primeiro foi notada por James Watson e Francis Crick quando formularam seu modelo da estrutura do DNA. A Figura 15.8b demonstra alguns mal pareamentos que resultam da mudança de um tautômero para outro, chamada de **mudança tautomérica**. A Figura 15.9 mostra como as mudanças tautoméricas afetam o resultado da replicação do DNA e resultam em prole mutante.

Os mal pareamentos também podem resultar quando uma das bases torna-se *ionizada*. Esse tipo de mal pareamento pode ocorrer mais frequentemente que os mal pareamentos devido a formas imino e enol das bases.

Todos os mal pareamentos descritos até agora levam a mutações de transição, nas quais uma purina é substituída por outra purina ou uma pirimidina por outra pirimidina (veja Figura 15.2). A DNA polimerase III bacteriana (veja o Capítulo 7) tem uma capacidade de edição que reconhece tais mal pareamentos e os remove, reduzindo muito o núme-

As bases podem ter formas tautoméricas propensas a mal pareamento

(a) Pareamento normal de base



(b) Bases mal pareadas

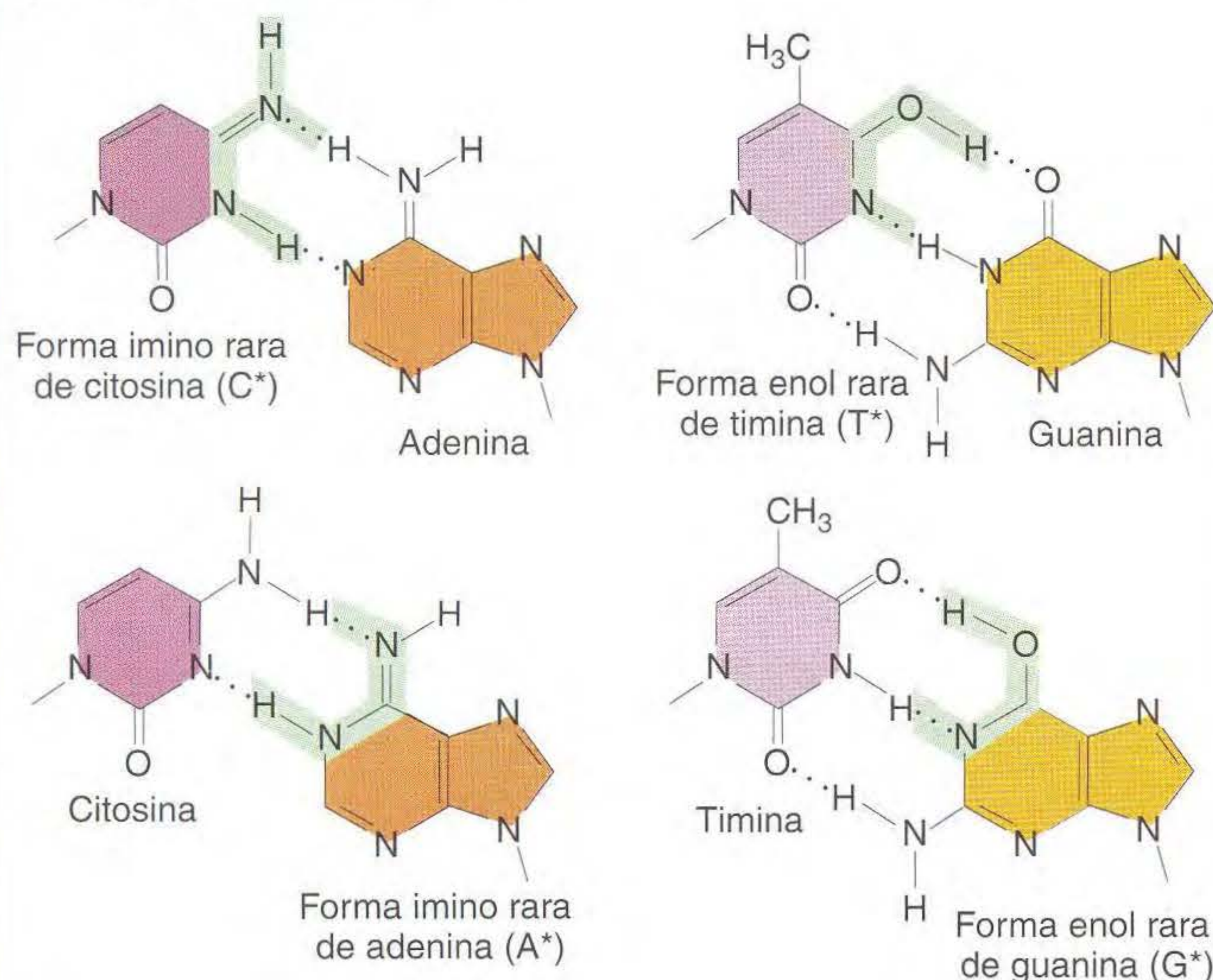


FIG. 15.8 Pareamento normal de bases comparado a bases mal pareadas. (a) Pareamento entre as formas normais (ceto) das bases. (b) Raras formas tautoméricas de bases resultam em mal pareamentos.

ro de mutações observadas. Outros sistemas de reparo (descritos mais adiante no capítulo) corrigem muitas das bases mal pareadas que escapam à correção pela função de edição da polimerase.

Transversões Nas mutações de transversão, uma pirimidina substitui uma purina ou vice-versa (veja Figura 15.2). As transversões não podem ser geradas pelos pareamentos errados mostrados na Figura 15.8. Com as bases do DNA na orientação normal, a criação de uma transversão por um erro de replicação iria requerer, em algum ponto no curso da replicação, o pareamento errado de uma purina com uma purina, ou uma pirimidina com uma pirimidina. Embora a dimensão da dupla hélice de DNA torne tais mal pareamentos energeticamente desfavoráveis, sabemos, dos estudos de difração de raios X, que os pares G-A, bem como outros pares de purina-purina, podem se formar.

Mudanças de matriz de leitura Os erros de replicação também podem levar a mudanças de matriz de leitura. Lembre do Capítulo 9 que tais mutações resultam em proteínas muito alteradas.

Embora alguns erros de replicação produzam mutações de substituição de bases, outros tipos de erros de replicação podem levar a **mutações indel**, isto é, inserções ou deleções de um ou mais pares de bases. Quando tais mutações adicionam ou subtraem um número de bases não-divisível por três (o tamanho de um códon), elas produzem mudanças de matriz de leitura nas regiões codificantes de proteína. O modelo prevalecente (Figura 15.10) sugere que as indels surgem quando são estabilizadas alças nas regiões unifilamentares por “pareamento deslocado (*slipped*)” de seqüências repetidas no curso da replicação. Esse mecanismo às vezes é chamado de *replicação deslocada*.

Mudanças tautoméricas nas bases do DNA podem causar mutações

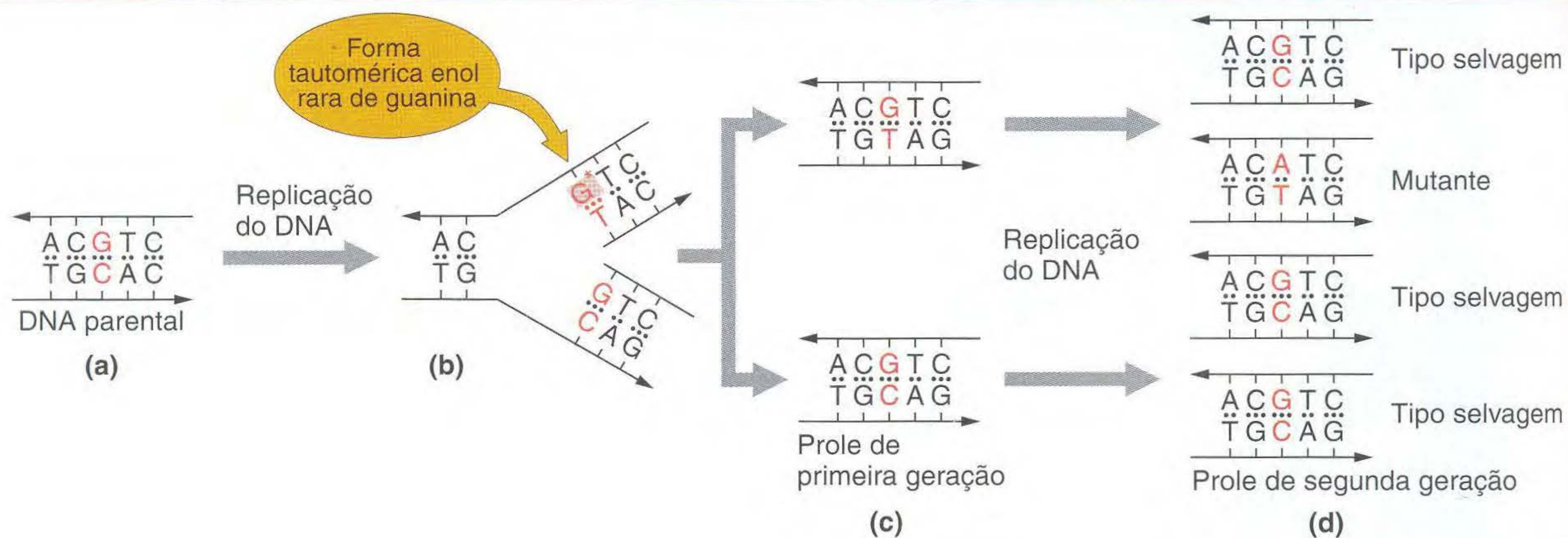


FIG. 15.9 Uma mudança tautomérica cria uma mutação em parte da prole após replicação do DNA. (a) No exemplo diagramado, uma guanina sofre uma mudança tautomérica para sua rara forma enol (G^*) na época da replicação. (b) Em sua forma enol, ela pareia com timina. (c e d) Na replicação seguinte, a guanina volta à sua forma cetona mais estável. A timina incorporada em oposição à forma enol de guanina, vista na parte b, dirige a incorporação de adenina na replicação subsequente, mostrada nas partes c e d. O resultado final é uma mutação $G \cdot C \rightarrow A \cdot T$. [De E. Gardner and D.P. Snustad, *Principles of Genetics*, 5th ed. (c) 1984 by John Wiley & Sons, New York.]

As mutações indel resultam em mudanças de matriz de leitura

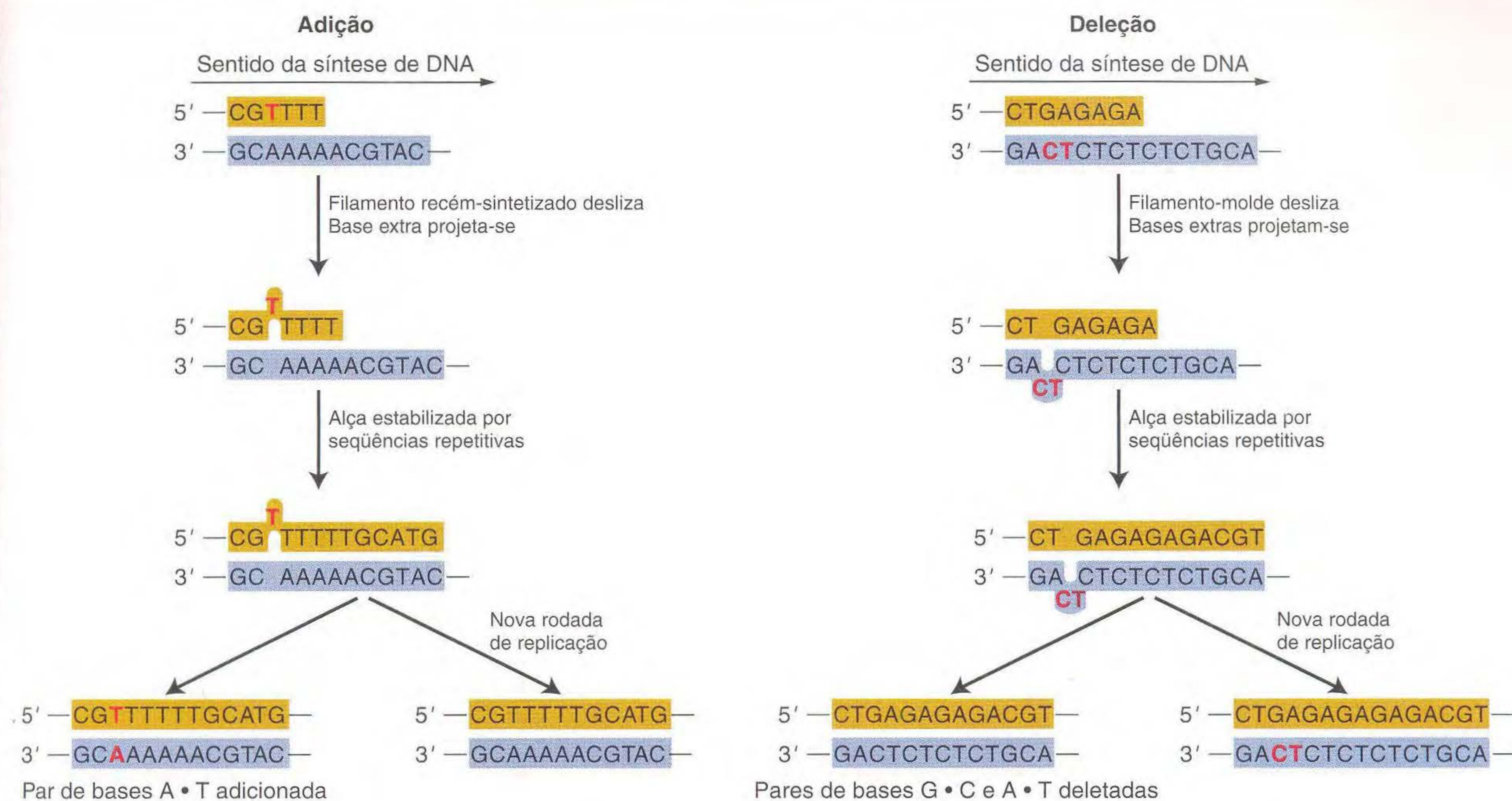
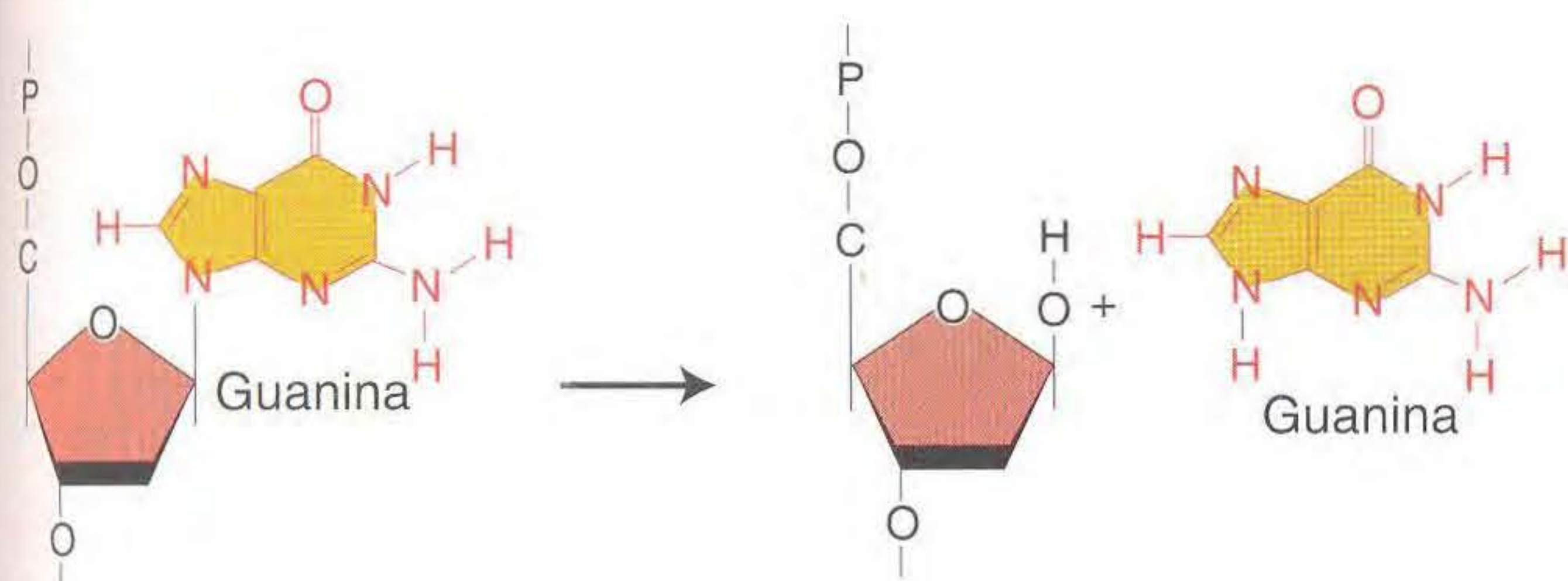


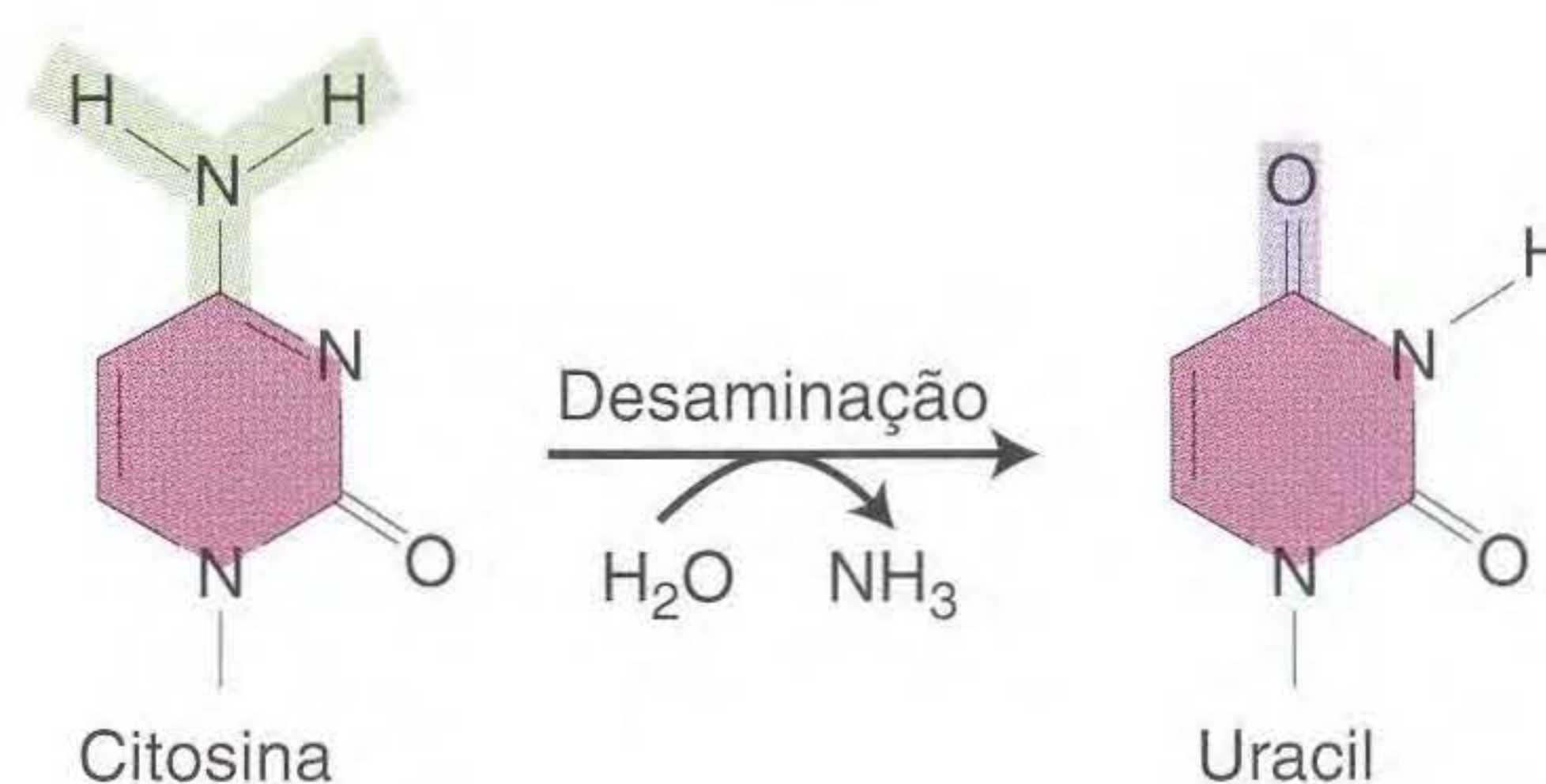
FIG. 15.10 Inserções de bases e deleções (mutações indel) causam mudanças de matriz de leitura pelo mal pareamento de seqüências repetidas no curso da replicação.

Lesões espontâneas Além dos erros de replicação, as **lesões espontâneas**, os danos de ocorrência natural ao DNA, podem gerar mutações. Duas das mais freqüentes lesões espontâneas resultam de depurinação e desaminação.

A depurinação, a mais comum das duas, é a perda de uma base purina. A depurinação consiste na interrupção da ligação glicosídica entre a base e a desoxirribose e a perda subsequente de uma guanina ou uma adenina do DNA.



A desaminação de citosina produz uracil.



Uracil não reparado irá parear com adenina na replicação, resultando na conversão de um par G · C em um par A · T (uma transição G · C → A · T).

As bases danificadas oxidativamente representam um terceiro tipo de lesão espontânea implicada na mutagênese. As espécies de oxigênio ativo, tais como radicais superóxido ($O_2^{\cdot-}$), peróxido de hidrogênio (H_2O_2) e radicais hidroxila ($\cdot OH$), são produzidos como subprodutos do metabolismo aeróbico normal. Eles podem causar dano oxidativo ao DNA, bem como a precursores do DNA (tais como GTP), resultando em mutação. As mutações de danos oxidativos foram implicadas em um número de doenças humanas. A Figura 15.11 mostra dois produtos de danos oxidativos. O produto 8-oxo-7-hidroxidesoxiguanosina (8-oxo dG ou GO) freqüentemente produz pareamentos errados com A, resultando em um alto nível de transversões G → T. O produto timidina glicol bloqueia a replicação do DNA se não for reparado, mas ainda não foi implicado na mutagênese.

Os radicais oxigênio podem danificar o DNA

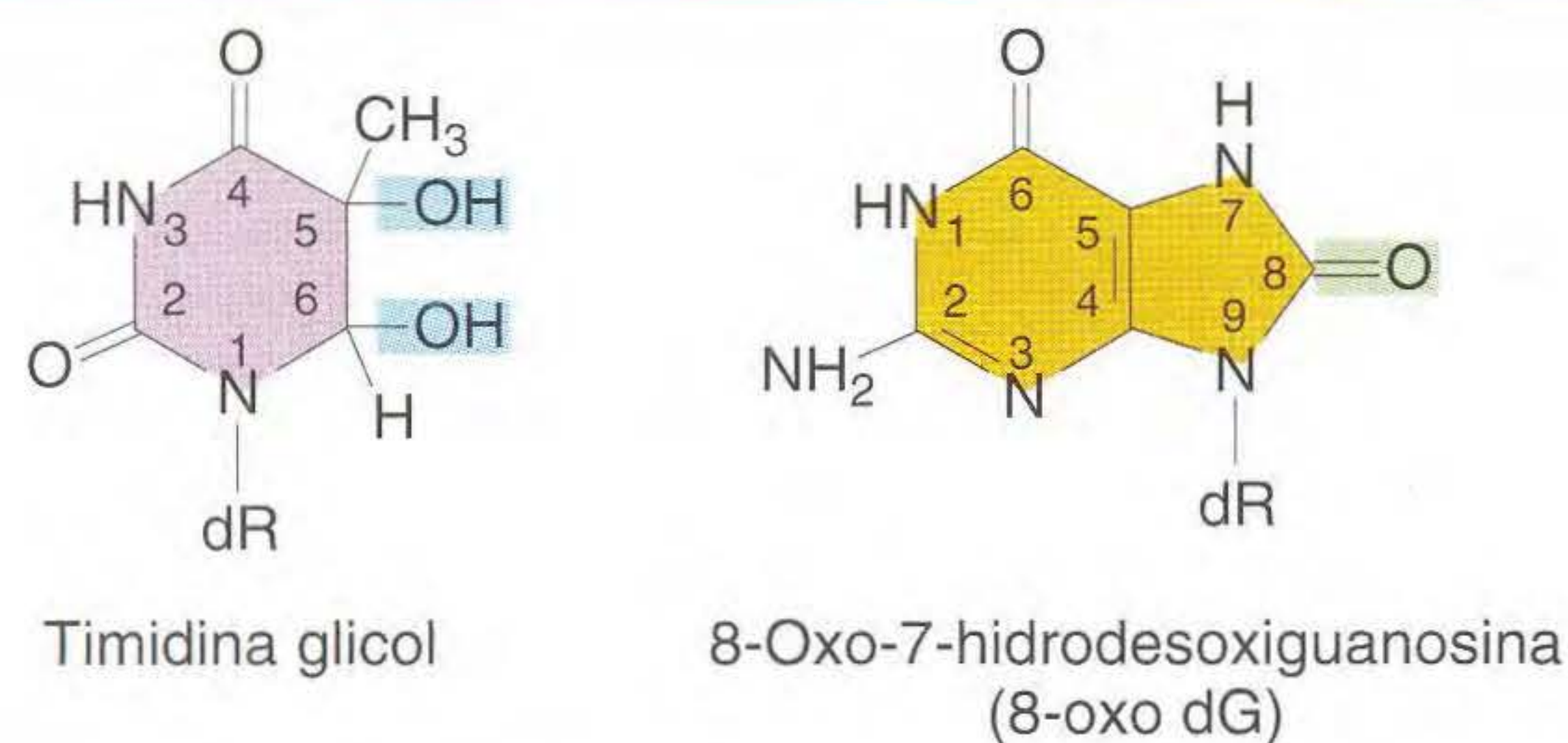


FIG. 15.11 Produtos formados após o DNA ter sido atacado por radicais oxigênio. Abreviação: dR, desoxirribose.

As mutações espontâneas podem ser geradas por processos diferentes. Os erros de replicação e as lesões espontâneas geram a maioria das substituições de bases espontâneas. Os erros de replicação podem causar também deleções que levam a mudanças de matriz de leitura.

Mutações espontâneas em humanos: doenças de repetições de trinucleotídeos

A análise de seqüências de DNA revelou que as mutações gênicas contribuem para numerosas doenças humanas hereditárias. Muitas são substituições de bases esperadas ou tipo indel de um só par de bases. Entretanto, algumas mutações são mais complexas. Vários desses distúrbios humanos são devidos a duplicações de seqüências curtas repetidas.

Um mecanismo comum responsável por várias doenças genéticas é a expansão de uma repetição de três pares de bases. Por esse motivo, elas são chamadas de doenças de **repetição de trinucleotídeos**. Um exemplo é uma doença humana chamada de *síndrome do X frágil*. Essa doença é a forma mais comum de prejuízo mental herdado, ocorrendo em cerca de 1 em 1.500 homens e 1 em 2.500 mulheres. Ela se manifesta citologicamente por um sítio frágil no cromossomo X que resulta em quebras *in vitro* (mas não leva ao fenótipo da doença). A síndrome do X frágil resulta de mudanças no número de uma repetição (CGG)_n em uma região do gene *FMR-1* que é transcrita mas não traduzida (Figura 15.12a).

Como o número de repetições está correlacionado ao fenótipo da doença? Os humanos normalmente mostram considerável variação no número de repetições CGG no gene *FMR-1*, variando de 6 a 54, com o alelo mais freqüente contendo 29 repetições. Às vezes, os genitores não-afetados e os avós dão origem a uma prole grande com síndrome do X frágil. A prole com os sintomas da doença tem enorme número de repetições, variando de 200 a 1.300 (Figura 15.12b). Os genitores não-afetados e os avós também foram vistos contendo número aumentado de cópias da repetição, mas variando de apenas 50 a 200. Por esse motivo, esses ancestrais foram ditos tendo *pré-mutações*. As repetições desses alelos de pré-mutação não são suficientes para causar o fenótipo da doença, mas são muito mais instáveis (isto é, facilmente expandidos) que os alelos normais, e assim eles levam a uma expansão maior

em sua prole. (Em geral, quanto mais expandida a repetição, maior parece ser a instabilidade.)

O mecanismo proposto para a geração dessas repetições é o deslizamento de mal pareamento no curso da síntese de DNA (Figura 15.13). Entretanto, a freqüência extraordinariamente alta de mutação nas repetições de trinucleotídeos na síndrome do X frágil sugere que, nas células humanas, após um limiar de cerca de 50 repetições a maquinaria de replicação não pode replicar fielmente a seqüência correta e ocorrem grandes variações no número de repetições.

Outras doenças, tais como a doença de Huntington (veja o Capítulo 2), também estão associadas à expansão de repetições de trinucleotídeos em um gene. Vários temas gerais aplicam-se a essas doenças. Na doença de Huntington, por exemplo, o gene tipo selvagem *HD* inclui uma seqüência repetida, em geral dentro da região codificante da proteína, e a mutação está correlacionada com uma considerável expansão dessa região repetida. A gravidade da doença está correlacionada ao número de cópias repetidas.

A doença de Huntington e a doença de Kennedy (também chamada de *atrofia muscular bulbar e espinhal ligada ao X*) resultam da amplificação de uma repetição de três pares de bases, CAG. As pessoas não-afetadas têm uma média de 19 a 21 repetições CAG, enquanto os pacientes afetados têm uma média de cerca de 46. Na doença de Kennedy, que é caracterizada por fraqueza muscular progressiva e atrofia, a expansão da repetição de trinucleotídeos é no gene que codifica o receptor de andrógeno.

As propriedades comuns a algumas doenças de repetição de trinucleotídeos sugerem um mecanismo comum pelo qual são produzidos fenótipos anormais. Primeiro, muitas dessas doenças parecem incluir neurodegeneração, isto é, morte celular dentro do sistema nervoso central. Segundo, em tais doenças as repetições de trinucleotídeos ocorrem nas matrizes de leitura abertas dos transcritos desses genes, levando a expansões ou contrações do número de repetições de um único aminoácido no polipeptídeo (por exemplo, as repetições CAG codificam uma repetição de poliglutamina). Assim, é fácil compreender por que essas doenças envolvem expansões de códon com três pares de bases.

Entretanto, essa explicação não é válida para todas as doenças de repetição de trinucleotídeos. Na síndrome do X frágil, a expansão de trinucleotídeos é perto da ponta 5' do mRNA de *FMR-1*, antes do sítio de início da tradução. Assim, não podemos atribuir as anomalias fenotípicas das mutações *FMR-1* a um efeito na estrutura da proteína. Uma chave para o problema com os genes mutantes *FMR-1* é que, ao contrário do gene normal, são hipermetiladas, uma característica associada aos genes transcricionalmente silenciados (veja Figura 15.12b). Com base nesses achados, a repetição expandida é suposta como levando a mudanças na estrutura da cromatina que silenciam a transcrição do gene mutante (veja o Capítulo 11). Em apoio a esse modelo está o achado de que o gene *FMR-1* é deletado em alguns pacientes com síndrome do X frágil. Essas observações apóiam uma mutação de perda de função.

As repetições de trinucleotídeos no gene *FMR-1* impedem transcrição

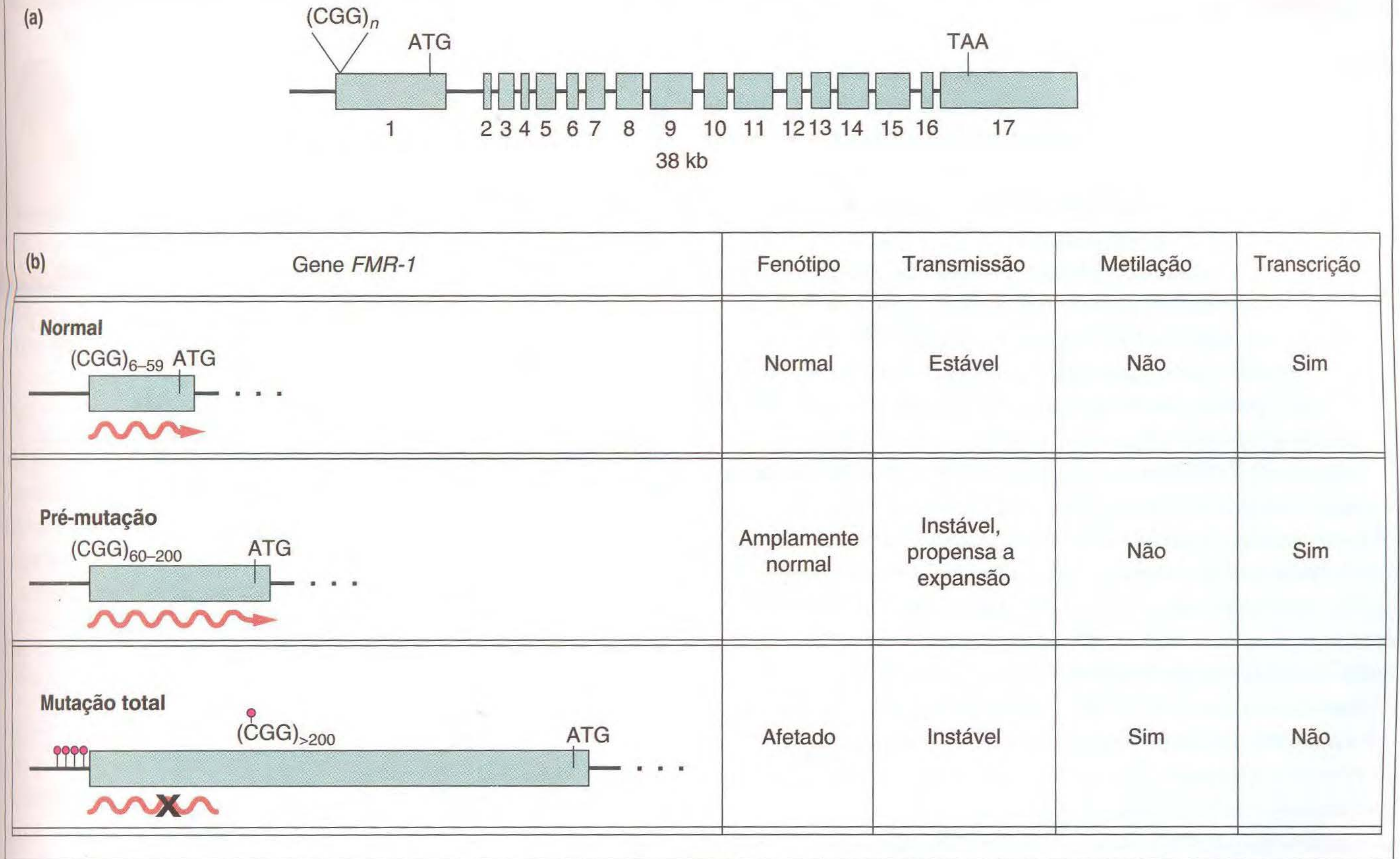


FIG. 15.12 O gene *FMR-1* na síndrome do X frágil. (a) Estrutura do éxon e repetição CGG antecedente. (b) Transcrição e metilação em alelos normais, pré-mutação e mutação total. Os círculos vermelhos representam grupos metil. [De W.T. O'Donnell and S.T. Warren, *Annu. Rev. Neurosci.* 25, 2002, 315-338, Fig. 1.]

Um mal pareamento deslocado causa expansão repetida

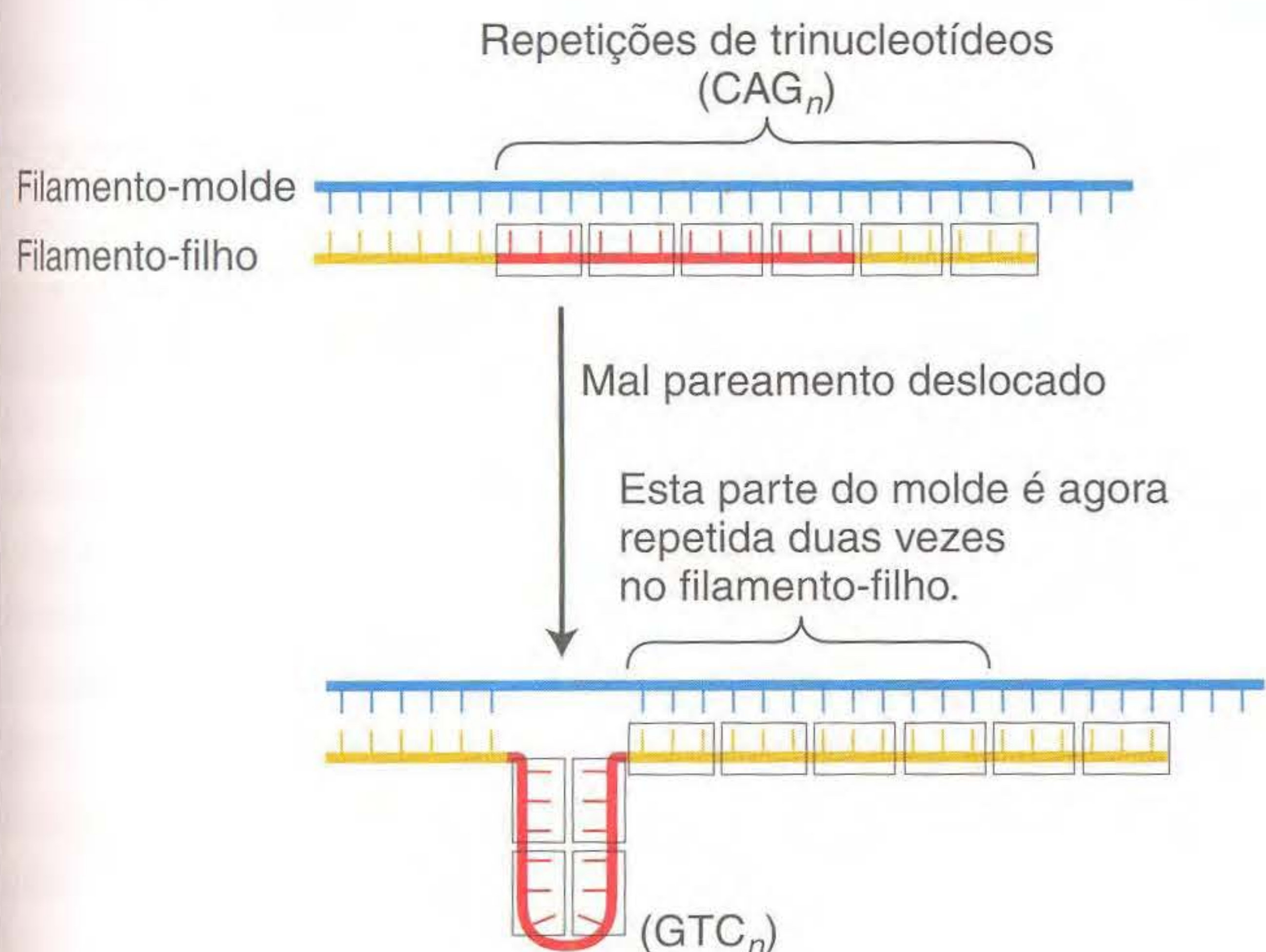


FIG. 15.13 Regiões de repetições de trinucleotídeos são propensas a mal pareamento (alça vermelha). Como consequência, a mesma região de repetição de trinucleotídeos pode ser duplicada duas vezes no curso da replicação.

As doenças de repetição de trinucleotídeos surgem pela expansão do número de cópias de uma sequência de três pares de bases presente em várias cópias, geralmente dentro da região codificante de um gene.

► 15.3 A Base Molecular das Mutações Induzidas

Enquanto algumas mutações são espontaneamente produzidas dentro da célula, outras fontes de mutação estão presentes no ambiente, seja intencionalmente aplicadas no laboratório ou encontradas por acaso no curso cotidiano. A produção das mutações no laboratório pela exposição a mutágenos é chamada de **mutagênese**, e o organismo é dito **mutagenizado**.

Mecanismos de mutagênese

Os mutágenos induzem mutações por pelo menos três mecanismos diferentes. Eles podem *substituir* uma base no DNA, *alterar* uma base de modo que ela faça um pareamento especificamente errado com outra base, ou *danificar* uma base de modo que ela não possa mais parear com qualquer base sob

condições normais. Mutar genes e observar as consequências fenotípicas é uma das estratégias primárias usadas pelos geneticistas.

Incorporação de análogos de bases Alguns compostos químicos são suficientemente similares às bases nitrogenadas normais do DNA de modo que sejam ocasionalmente incorporadas ao DNA em lugar das bases normais; tais compostos são chamados de **análogos de bases**. Após estarem no lugar, esses análogos têm propriedades de pareamento diferentes das bases normais; assim, podem produzir mutações fazendo com que nucleotídeos incorretos sejam inseridos em oposição a eles na replicação. O análogo original de bases existe apenas em um único filamento, mas pode causar uma substituição de par de nucleotídeos que é replicada em todas as cópias do DNA descendentes do filamento original.

Por exemplo, 5-bromouracil (5-BU) é um análogo de timina que tem bromo na posição carbono-5 em lugar do grupo CH_3 encontrado na timina. Essa mudança não afeta os átomos que tomam parte nas pontes de hidrogênio no pareamento de bases, mas a presença do bromo altera significativamente a distribuição de elétrons na base. A estrutura normal (a forma ceto) de 5-BU faz par com adenina, como mostrado na Figura 15.14a. Entretanto, 5-BU pode frequentemente mudar para a forma enol ou uma forma ionizada; esta última pareando *in vivo* com guanina (Figura 15.14b). Assim, a identidade do par formado na replicação irá depender da forma de 5-BU no momento do pareamento. O 5-bromouracil causa transições quase exclusivamente, como mostrado na Figura 15.14.

Outro análogo amplamente usado em pesquisa é 2-aminopurina (2-AP). Esse análogo de adenina pode parear com timina, mas também pode fazer um mal pareamento com citosina quando protonado, como mostrado na Figura 15.15. Portanto, quando 2-AP é incorporado ao DNA por pareamento com timina, ele pode gerar transições $\text{A} \cdot \text{T} \rightarrow \text{G} \cdot \text{C}$ por mal pareamento com citosina nas replicações subsequentes. Ou, se 2-AP é incorporado por mal pareamento com citosina, então as transições $\text{G} \cdot \text{C} \rightarrow \text{A} \cdot \text{T}$ resultarão quando ela pareia com timina. Os estudos genéticos mostraram que 2-AP, como 5-BU, causam transições quase exclusivamente.

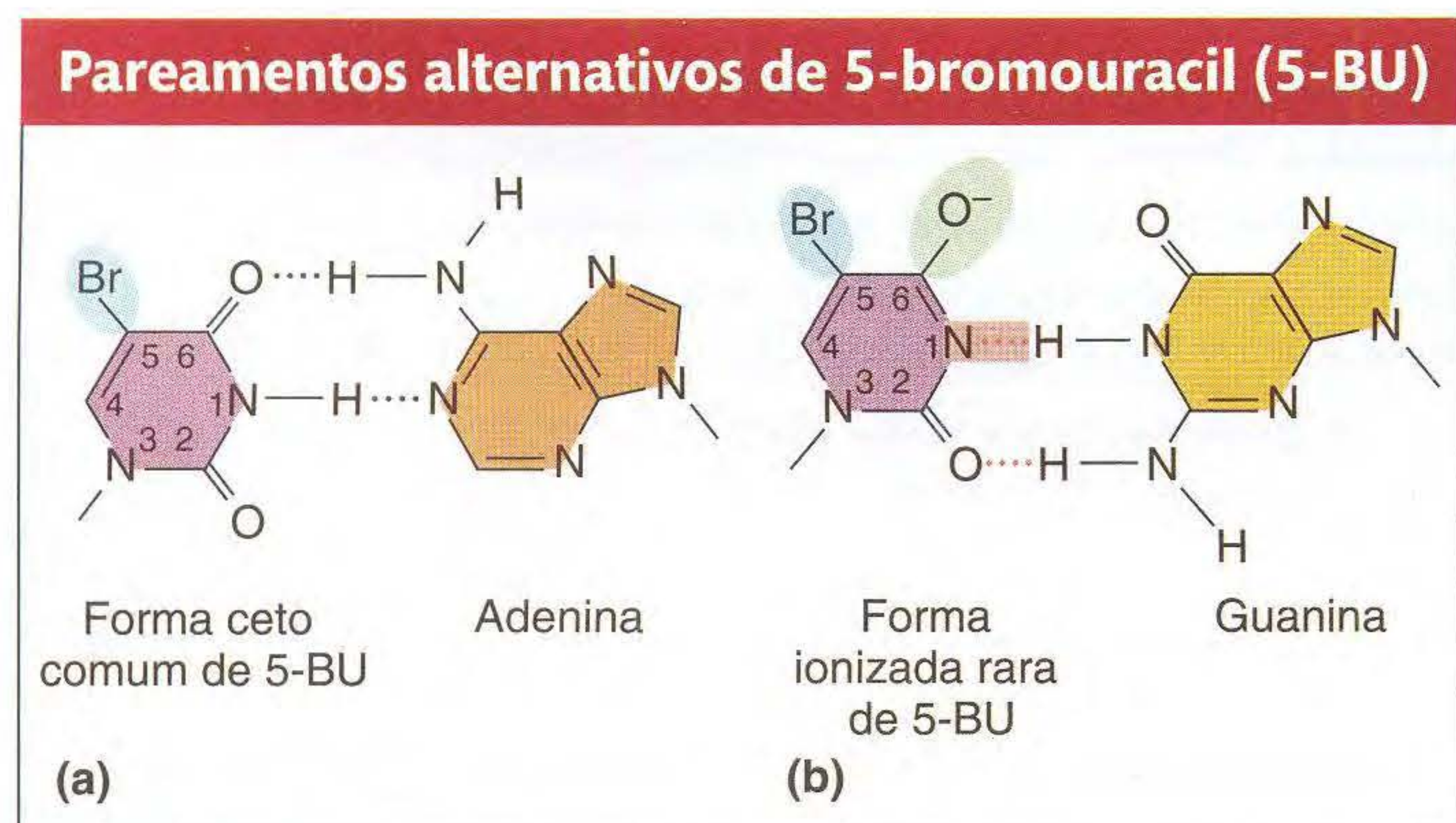


FIG. 15.14 (a) Um análogo de timina, 5-BU pode ser erradamente incorporado ao DNA como uma base. (b) A forma ionizada do pareamento de base com guanina.

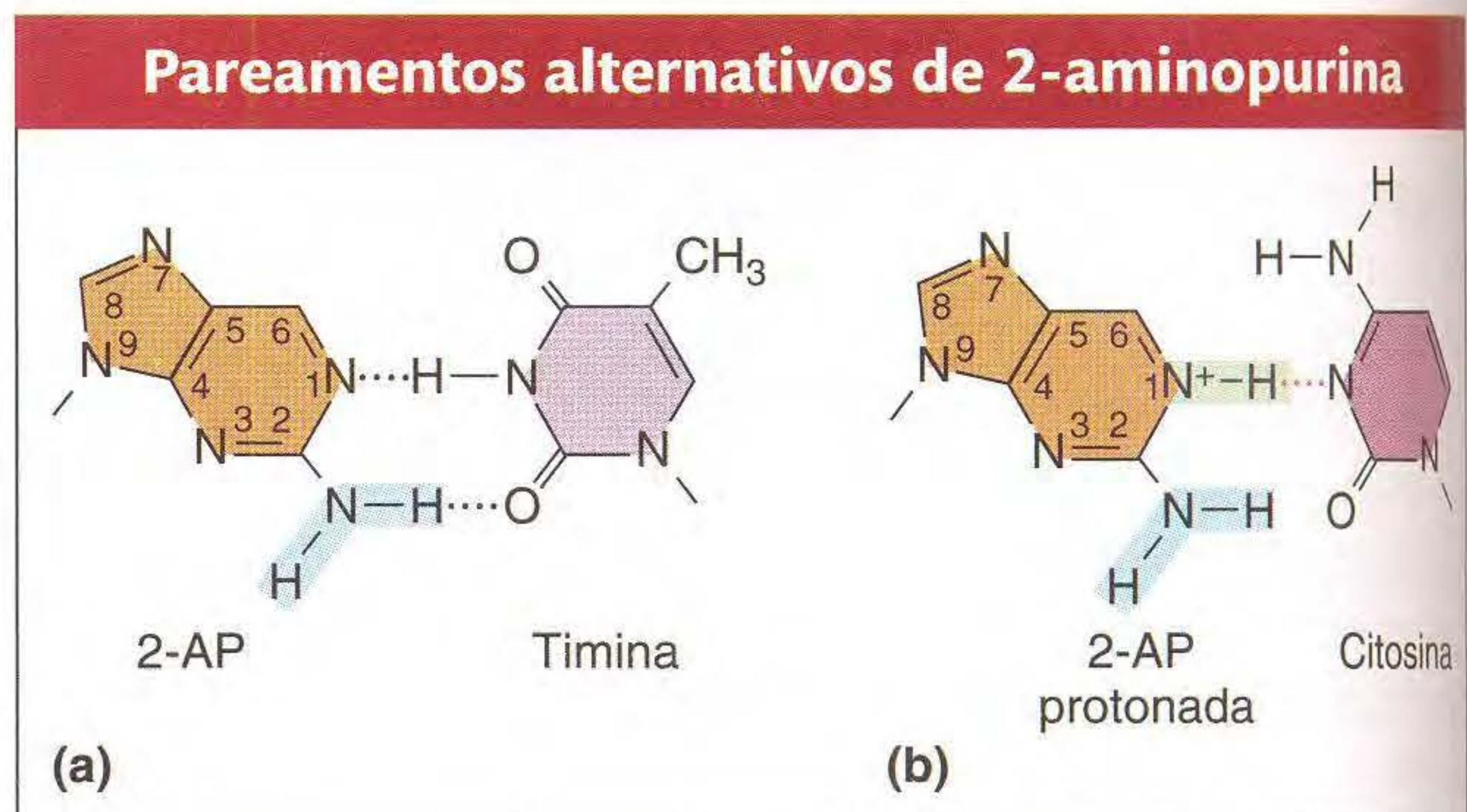
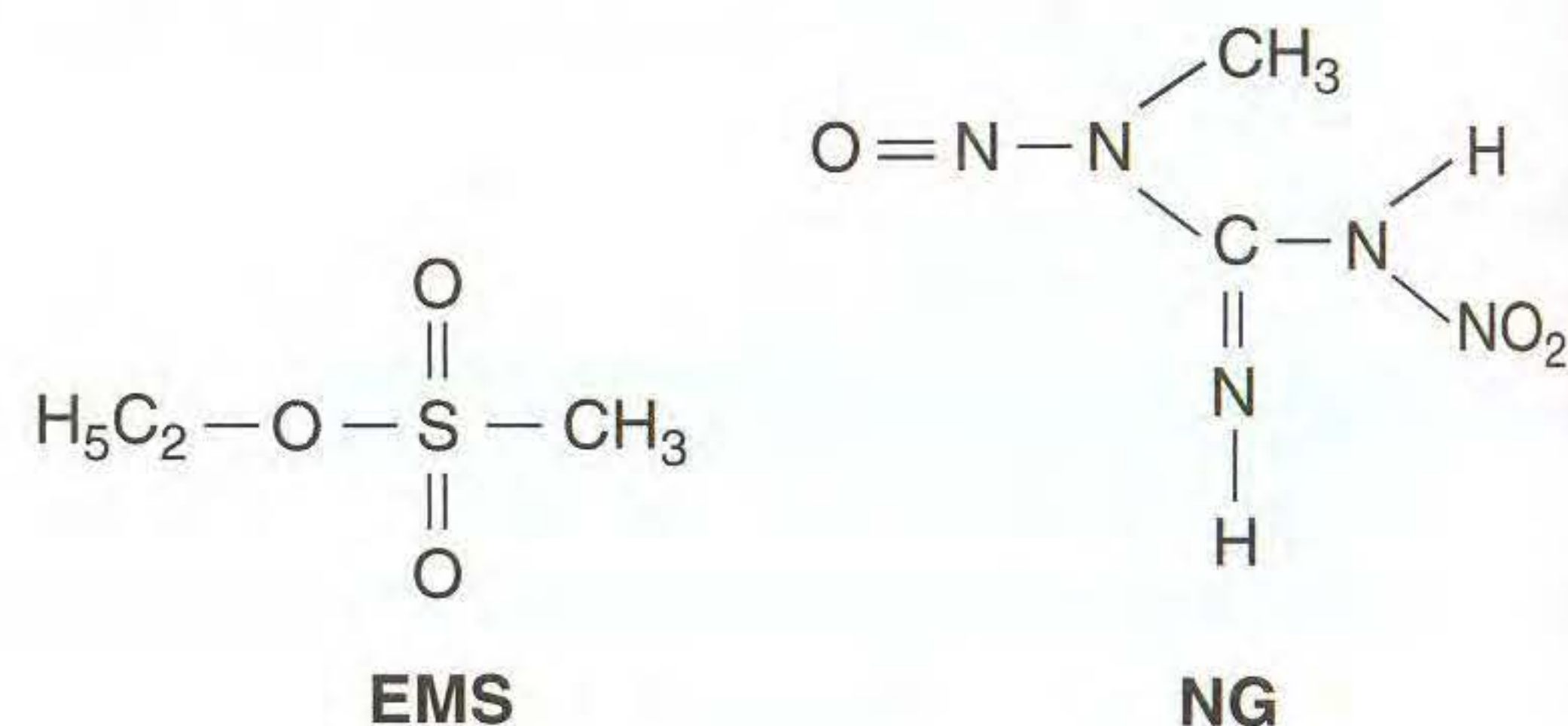


FIG. 15.15 (a) Um análogo de adenina, 2-aminopurina (2-AP) pode parear com timina. (b) Em seu estado protonado, 2-AP pode parear com citosina.

Mal pareamento específico Alguns mutágenos não são incorporados no DNA, mas alteram uma base de tal modo que se forma um mal pareamento específico. Alguns agentes alquilantes, tais como o etilmetanossulfato (EMS) e a amplamente usada nitrosoguanidina (NG), operam por essa via.



Tais agentes adicionam grupos alquila (um grupo etila em EMS e um grupo metila em NG) para muitas posições em todas as quatro bases. Entretanto, a formação de uma mutação é mais bem correlacionada com a adição de oxigênio na posição 6 de guanina para criar uma *O*-6-alcilguanina. Essa adição leva a um mal pareamento direto com timina, como mostrado na Figura 15.16, e resultaria em transições $\text{G} \cdot \text{C} \rightarrow \text{A} \cdot \text{T}$ e próxima rodada de replicação.

Os agentes alquilantes também podem modificar as bases em dNTP (onde N é qualquer base), que são precursores na síntese de DNA.

Agentes intercalares Os **agentes intercalares** formam uma outra classe importante de modificadores de DNA. Esse grupo de compostos inclui proflavina, acridina laranja e uma classe de substâncias chamadas compostos ICR (Figura 15.17a). Esses agentes são moléculas planares que mimetizam pares de bases e são capazes de inserir-se (intercalarse) entre bases nitrogenadas empilhadas no cerne da dupla hélice de DNA (Figura 15.17b). Nessa posição intercalada, tal agente pode causar uma inserção ou deleção de um único par de nucleotídeos.

Danos a bases Um grande número de mutágenos danificam uma ou mais bases, e, assim, não é possível um pareamento específico de bases. O resultado é um bloqueio de replicação, porque a síntese de DNA não ocorrerá além de uma base que

Mal pareamentos específicos induzidos por alquilação

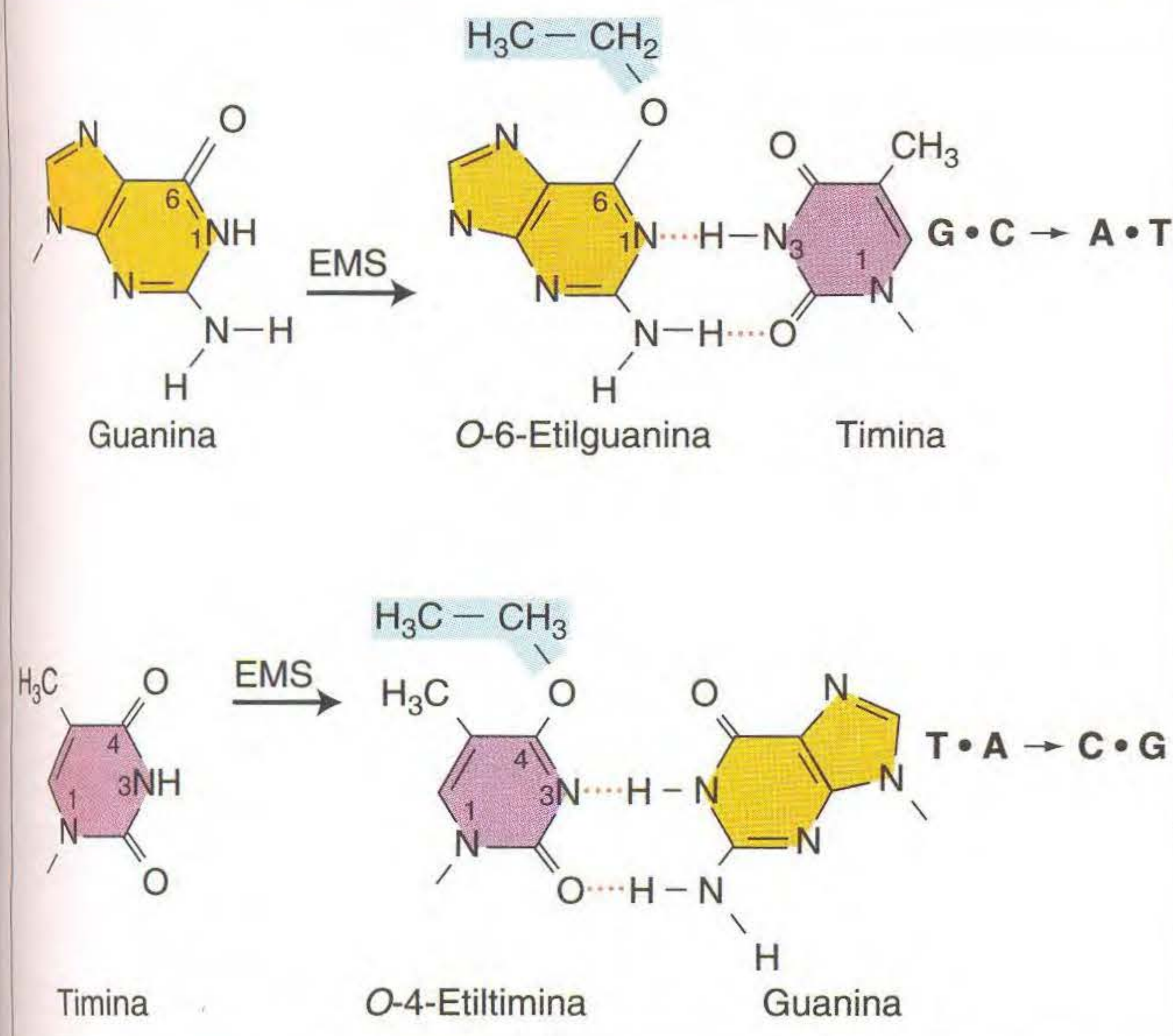


FIG. 15.16 Tratamento com EMS altera a estrutura da guanina e timina e leva a mal pareamentos.

não possa especificar seu parceiro complementar por pontes de hidrogênio. Os bloqueios de replicação podem causar mais mutação, como será explicado mais adiante no capítulo (veja a seção de reparo por excisão de nucleotídeo).

A luz ultravioleta geralmente causa danos a bases nucleotídicas na maioria dos organismos. A luz ultravioleta gera um número de tipos distintos de alterações no DNA, chamados de *fotoprodutos* (da palavra *foto* para “luz”). O mais provável desses produtos para levar a mutações são duas lesões diferentes que unem pirimidinas adjacentes no mesmo filamento. Essas lesões são o fotodímero de pirimidina ciclobutano e o fotoproduto 6-4 (Figura 15.18).

A radiação ionizante resulta na formação de moléculas ionizadas e excitadas que podem danificar o DNA. Devido à natureza aquosa dos sistemas biológicos, as moléculas gera-

das pelos efeitos da radiação ionizante na água produzem os maiores danos. Muitos tipos diferentes de espécies reativas de oxigênio são produzidos, mas os que mais danificam as bases do DNA são $\cdot\text{OH}$, O_2^- e H_2O_2 . Essas espécies levam à formação de adições diferentes e produtos de degradação. Entre os mais prevalentes, mostrados na Figura 15.11, estão timidina glicol e 8-oxo dG, ambos os quais podem resultar em mutações.

A radiação ionizante também pode danificar diretamente o DNA em vez de espécies de oxigênio reativo. Tal radiação pode causar a quebra de ligações N-glicosídicas, levando à formação de sítios apurínicos ou apirimidínicos, e pode causar quebras filamentosas. De fato, as quebras filamentosas são responsáveis pela maioria dos efeitos letais da radiação ionizante.



A aflatoxina B₁ é um poderoso carcinógeno que se liga à guanina na posição N-7 (Figura 15.19). A formação desse produto de adição leva à quebra da ligação entre a base e o açúcar, liberando assim a base e gerando um sítio apurínico. A aflatoxina B₁ é um membro de uma classe de carcinógenos químicos conhecidos como produtos de adição quando se ligam covalentemente ao DNA. Outros exemplos incluem o diol epóxido de benzo(a)pireno, um composto produzido por máquinas de combustão interna. Todos os compostos dessa classe induzem a mutações, embora por mecanismos que nem sempre estão claros.

Os mutágenos induzem mutações por uma variedade de mecanismos. Alguns mutágenos mimetizam bases normais e são incorporados ao DNA, onde podem parear errado. Outros danificam bases e causam mal pareamento específico ou destroem o pareamento causando o não-reconhecimento de bases.

O teste de Ames: avaliação de mutágenos em nosso ambiente



Um grande número de compostos químicos foi sintetizado, e muitos têm possíveis aplicações comerciais. Aprendemos arduamente que os benefícios poten-

Agentes intercalares

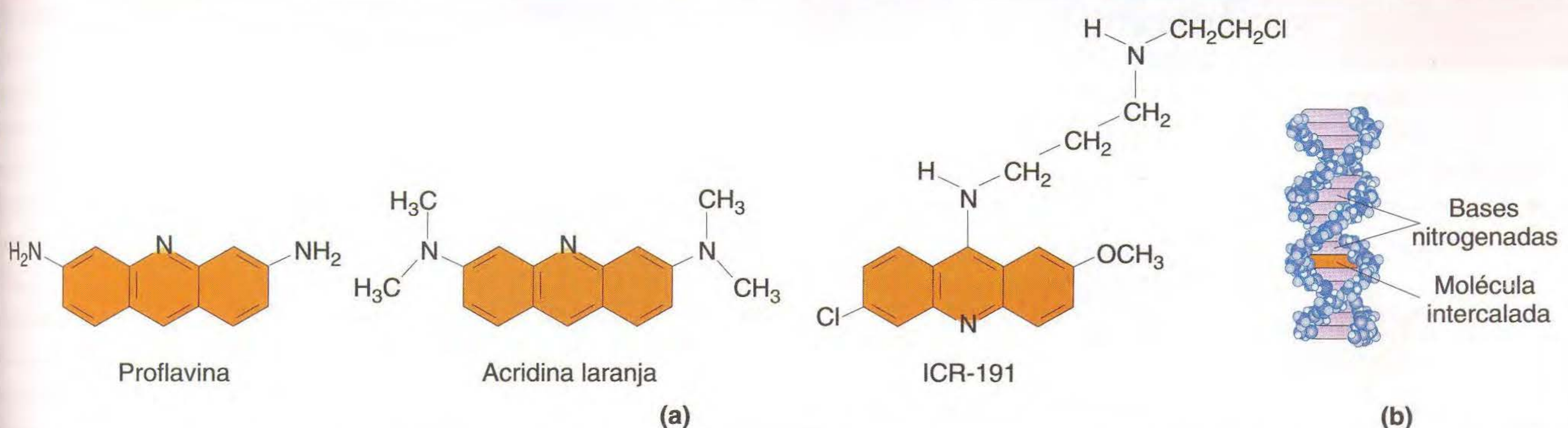


FIG. 15.17 Estrutura dos agentes intercalares comuns (a) e sua interação com o DNA (b). [De L.S. Lerman, “The Structure of the DNA-Acridine Complex,” *Proc. Natl. Acad. Sci. USA* 49, 1963, 94.]

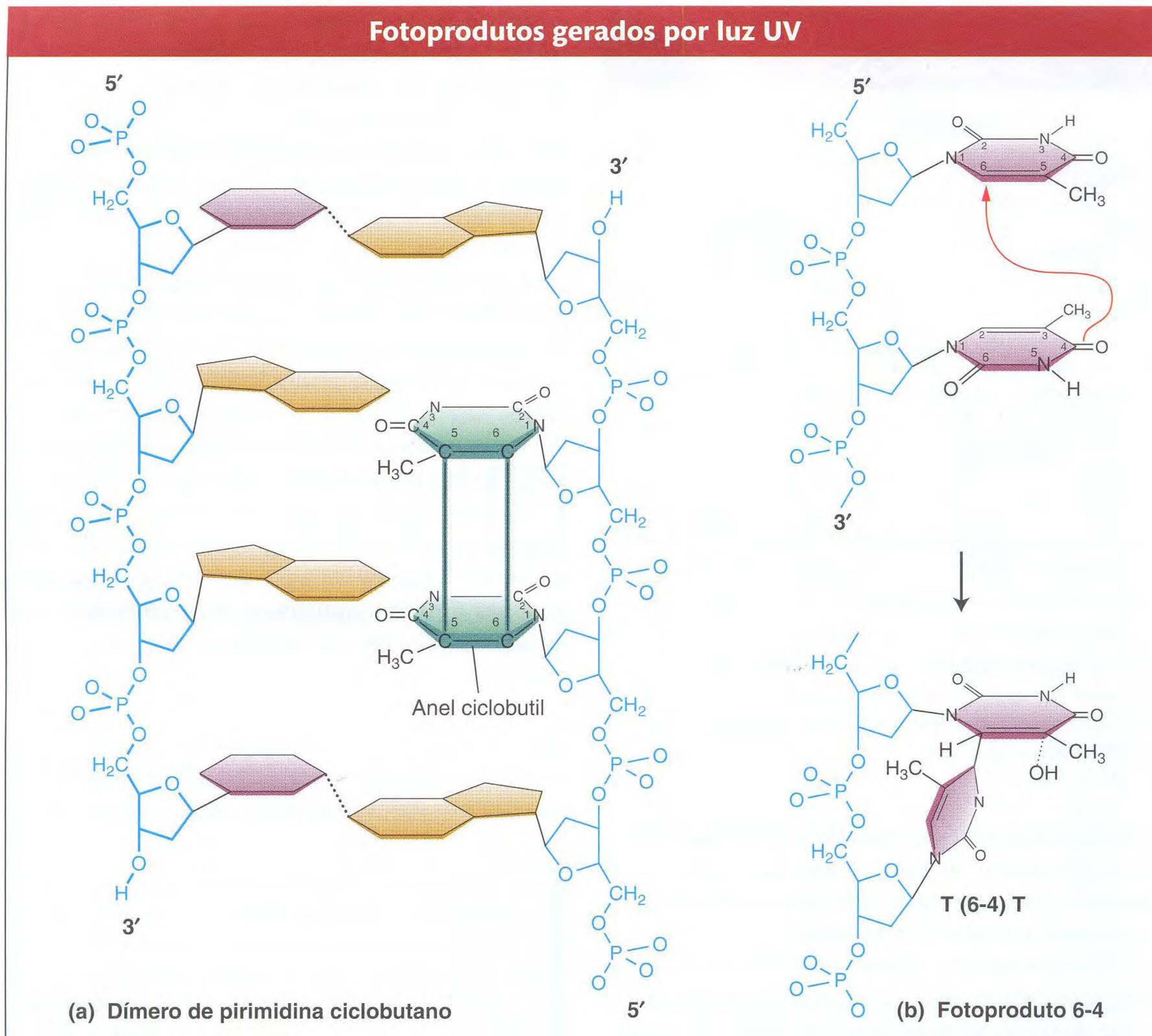


FIG. 15.18 Fotoprodutos que unem pirimidinas adjacentes no DNA estão fortemente correlacionadas a mutagênese. [(Esquerda) De E.C. Friedberg, *DNA Repair*. Copyright 1985 by W.H. Freeman and Company. (Direita) De J.S. Taylor et al.]

ciais dessas aplicações têm que ser avaliados em relação aos riscos de saúde e ambientais. Assim, é essencial ter técnicas eficientes de triagem para avaliar alguns dos riscos de um grande número de compostos.

Um importante fator de risco de muitos compostos é como agentes causadores de câncer (carcinógenos). Assim, é muito

importante ter sistemas-modelo válidos nos quais a carcinogênese dos compostos pode ser eficiente e efetivamente avaliada. Entretanto, usar um sistema-modelo de mamífero tal camundongo é muito lento, demorado e caro.

Nos anos de 1970, Bruce Ames reconheceu que existe uma forte correlação entre a carcinogênese e a mutagênese dos compostos. Ele percebeu que a medida das taxas de mutação nos sistemas bacterianos seria um modelo efetivo para avaliar a mutagenicidade de compostos como um primeiro nível de detecção de carcinógenos potenciais. Entretanto, ficou claro que nem todos os carcinógenos eram mutagênicos; alguns metabólitos carcinógenos produzidos no corpo são de fato agentes mutagênicos. Tipicamente, esses metabólitos são produzidos no fígado, e as reações enzimáticas que convertem os carcinógenos em metabólitos bioativos não ocorrem em bactérias.

Ames percebeu que podia superar esse problema tratando linhagens especiais da bactéria *Salmonella typhimurium* com extratos de fígados de rato contendo enzimas metabólicas (Figura 15.20). As linhagens especiais de *S. typhimurium* tinham um dos vários alelos mutantes de um gene responsável pela síntese de histidina que foram conhecidos como

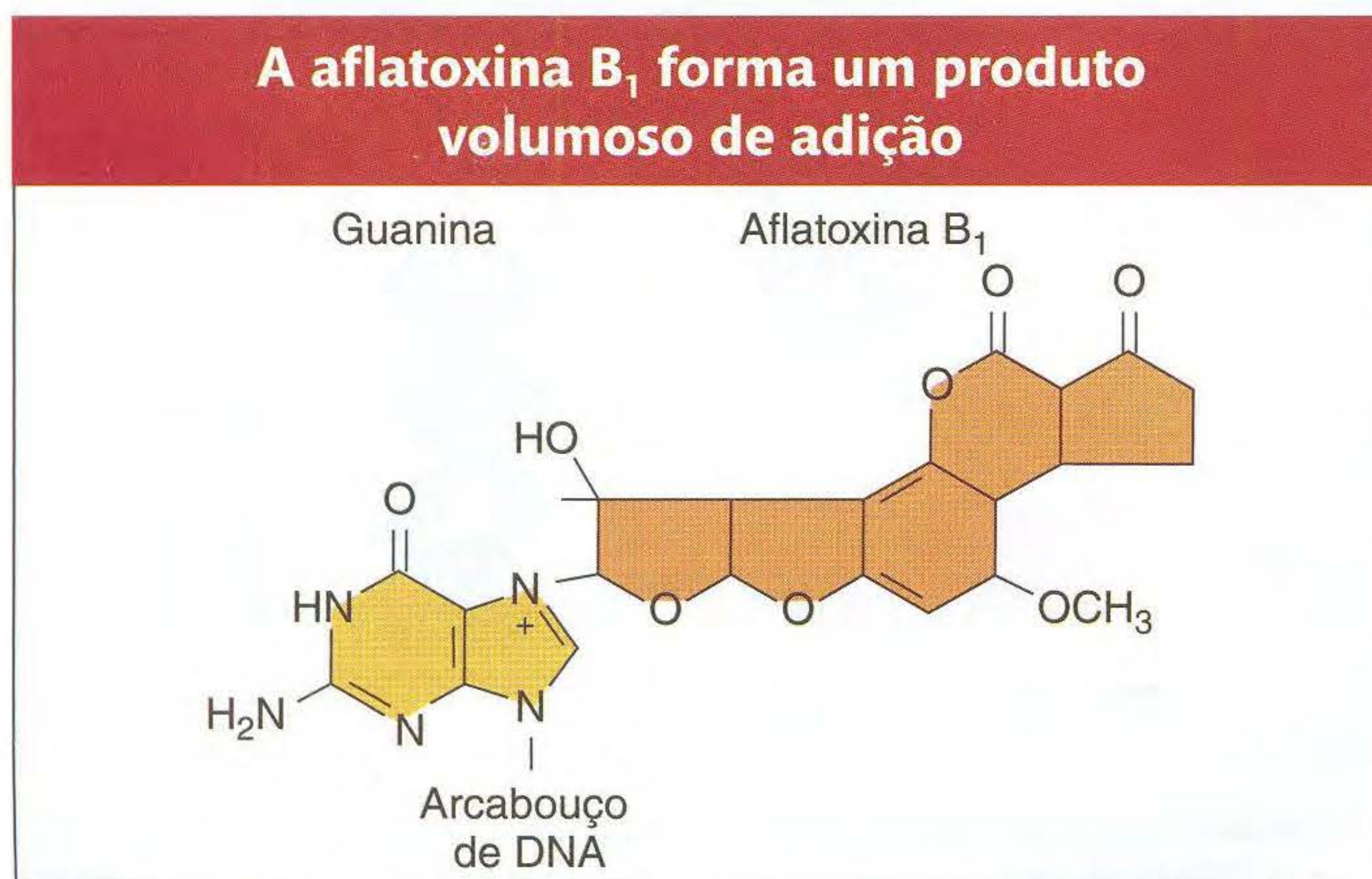


FIG. 15.19 Aflatoxina B₁ metabolicamente se ativada liga-se ao DNA.

O teste de Ames revela compostos mutagênicos

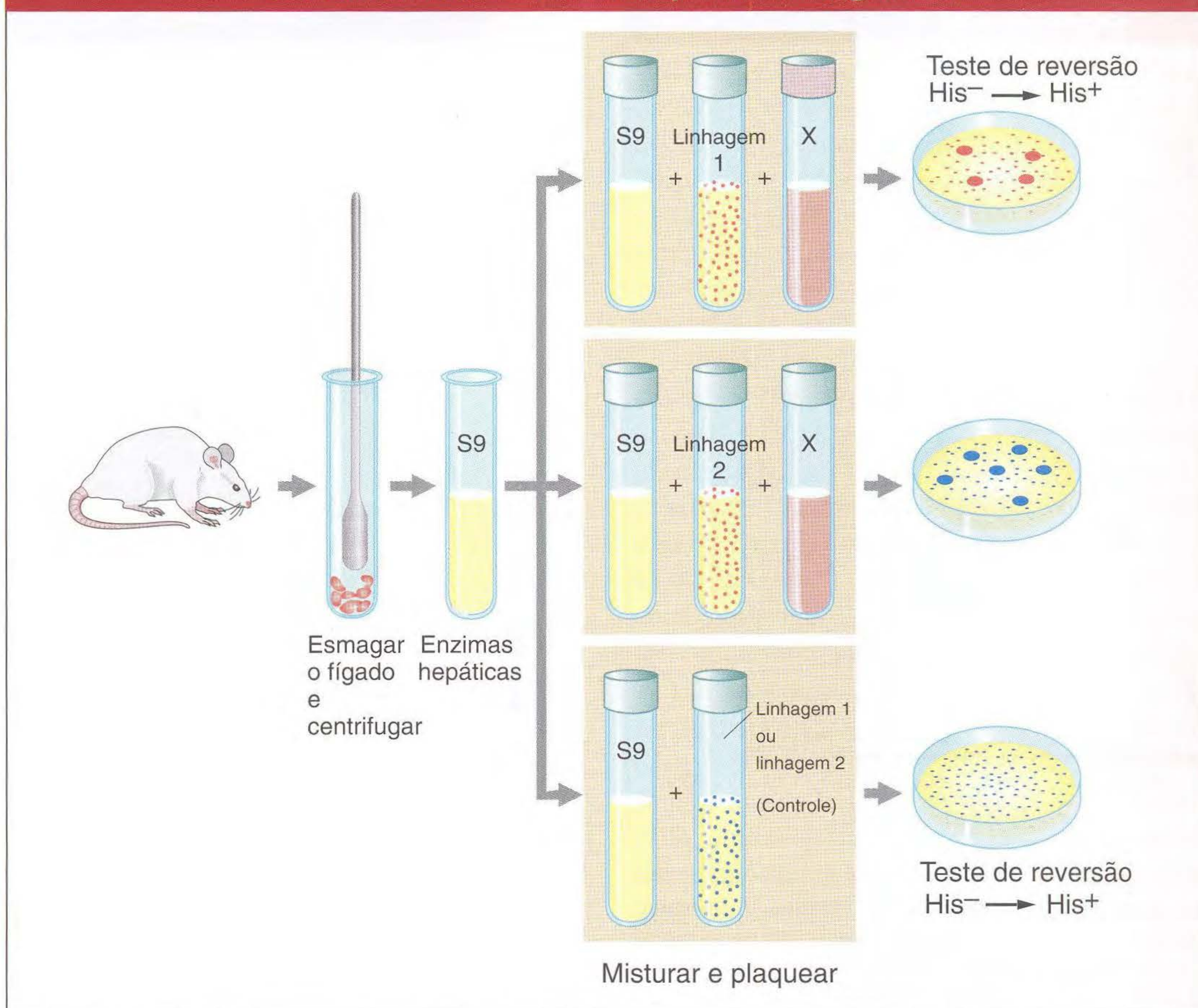


FIG. 15.20 Resumo dos procedimentos usados no teste de Ames. Enzimas hepáticas solubilizadas (S9) são adicionadas a uma suspensão de bactérias auxotróficas em uma solução do carcinógeno potencial (X). A mistura é plaqueada em um meio que não contém histidina. A presença de revertentes indica que a substância é um mutágeno e, possivelmente, também um carcinógeno.

“reverter” (isto é, retornar ao fenótipo tipo selvagem) apenas com alguns tipos de eventos mutacionais adicionais. Por exemplo, um alelo chamado TA100 pode ser revertido ao tipo selvagem apenas por mutação de substituição de bases, enquanto TA1538 e 1535 podiam ser revertidos apenas por mutações indel resultando em uma mudança de proteína (Figura 15.21).

As bactérias tratadas de cada uma dessas linhagens foram expostas ao composto do teste, e então cultivadas em placas de Petri em um meio sem histidina. A falta desse nutriente garantiu que apenas indivíduos revertentes que continham a substituição apropriada de base ou mudança de matriz de leitura iriam crescer. O número de colônias em cada placa e o número total de bactérias testadas foram determinados, permitindo a Ames medir a frequência da reversão. Os compostos que produziram metabólitos induzindo níveis elevados de reversão relativos a extratos de fígado controle não tratados seriam então claramente mutagênicos e, possivelmente, carcinógenos. O teste de Ames demonstrou assim ser um meio importante para triar milhares de compostos e avaliar um aspecto de seu risco para a saúde no ambiente. Esse teste ainda é usado hoje em dia como uma ferramenta importante para a avaliação da segurança de compostos químicos.

► 15.4 Mecanismos Biológicos de Reparo

Após o levantamento dos vários modos pelos quais o DNA pode ser danificado, a partir de fontes tanto de dentro da célula (replicação, oxigênio reativo, e assim por diante) como externas (ambientais: luz UV, radiação ionizante, mutágenos), você deve estar imaginando como a vida conseguiu sobreviver e se desenvolver por bilhões de anos. O fato é que os organismos variando de bactérias a humanos ou plantas podem reparar eficientemente seus DNA por uma variedade de mecanismos que, juntos, empregam cerca de até 100 proteínas de reparo conhecidas. De fato, nossa atual compreensão é que o DNA é a única molécula que os organismos reparam em vez de substituir. Como você verá, a falha desses sistemas de reparo é uma causa significativa de uma variedade de doenças humanas herdadas.

O mecanismo de reparo mais importante foi resumidamente mencionado no Capítulo 7, a função de revisão das DNA polimerases que replicam o DNA como parte do replissomo. Como notado no referido capítulo, tanto a DNA polimerase I quanto a DNA polimerase III são capazes de remover bases mal pareadas que foram inseridas erradamen-

Teste de Ames da mutagenicidade de aflatoxina B₁

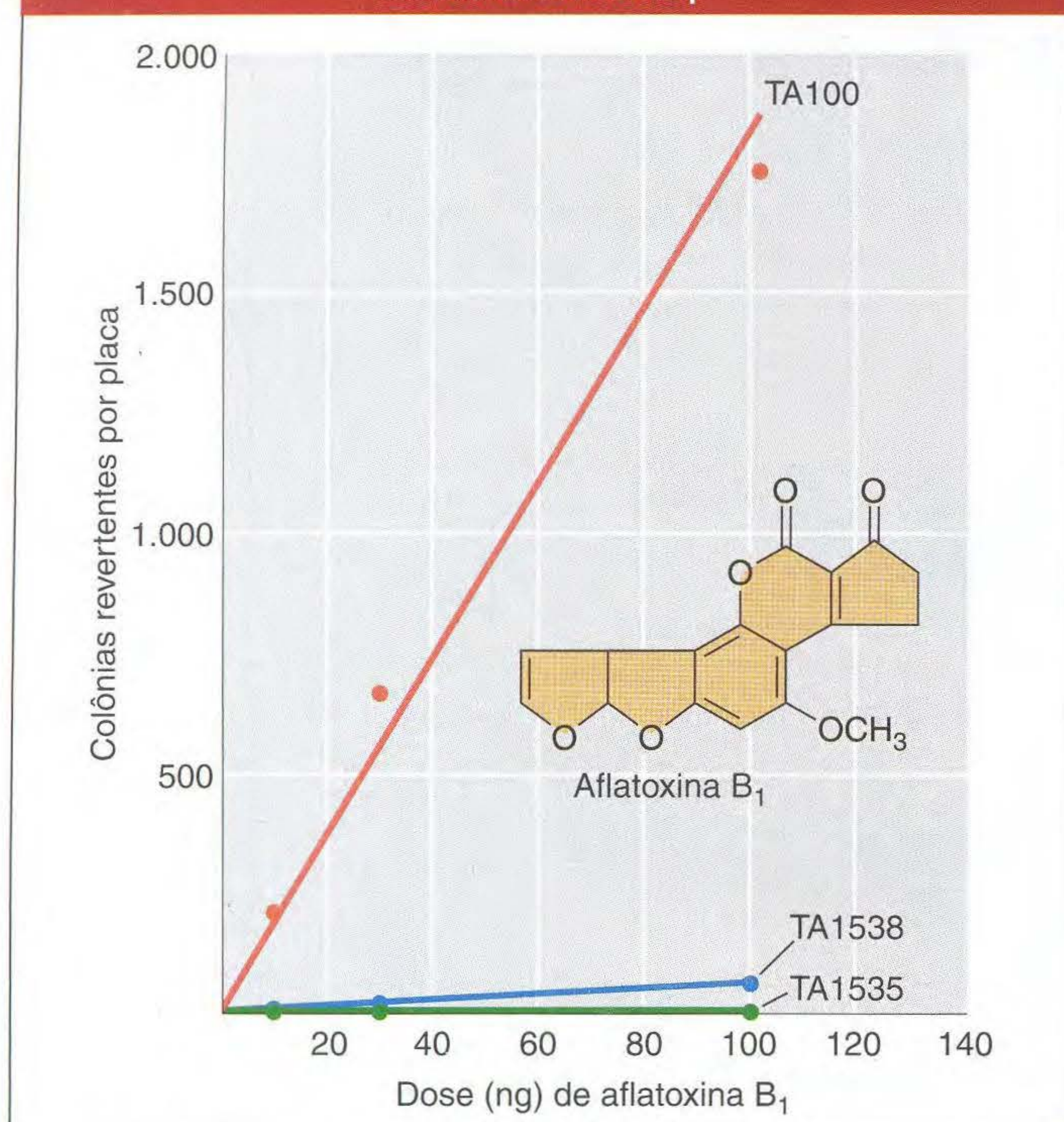


FIG. 15.21 TA100, TA1538 e TA1535 são linhagens de *Salmonella* portadoras de diferentes mutações auxotróficas de histidina. A linhagem TA100 é altamente sensível a reversão pela substituição do pareamento de bases. As linhagens TA1535 e TA1538 são sensíveis a reversão por meio de mudança de matriz de leitura. Os resultados do teste mostram que a aflatoxina B₁ é um potente mutágeno que faz substituições de pares de bases, mas não mudanças de matriz de leitura. [De]. McCann and B.N. Ames, in W.G. Flamm and M.A. Mehlman, eds. *Advances in Modern Technology*, vol. 5. Copyright by Hemisphere Publishing Corporation, Washington, D.C.]

te. Vamos examinar agora algumas das outras vias de reparo, começando com o reparo livre de erros.

Reversão direta de DNA danificado

O modo mais direto de reparar uma lesão é revertê-la diretamente, regenerando assim a base normal (Figura 15.22). Embora a maioria dos tipos de danos sejam essencialmente irreversíveis, as lesões podem ser reparadas por reversão direta em alguns casos. Um caso é um fotodímero mutagênico causado por luz UV. O dímero de pirimidina ciclobutano (CPD) pode ser reparado por uma enzima chamada de *CPD fotoliase*. A enzima liga-se ao fotodímero e o divide para regenerar as bases originais. Esse mecanismo de reparo é chamado de *fotorreativação* porque a enzima requer luz para funcionar. Outras vias de reparo são necessárias para remover danos de UV na ausência de luz do comprimento de onda apropriado (> 300 nm).

As alquiltransferases são enzimas que também reverterem diretamente as lesões. Elas removem alguns grupos alquila que foram adicionados à posição O-6 da guanina (veja Figura 15.16) por mutágenos, tais como nitrosoguanidina e etilmetanossulfonato. A metiltransferase de *E. coli* foi bem estudada.

Fotodímeros induzidos por UV e reparo por excisão

W.W. ANIMAÇÃO COMPUTADORIZADA

Um fotodímero pode ser revertido por reparo direto

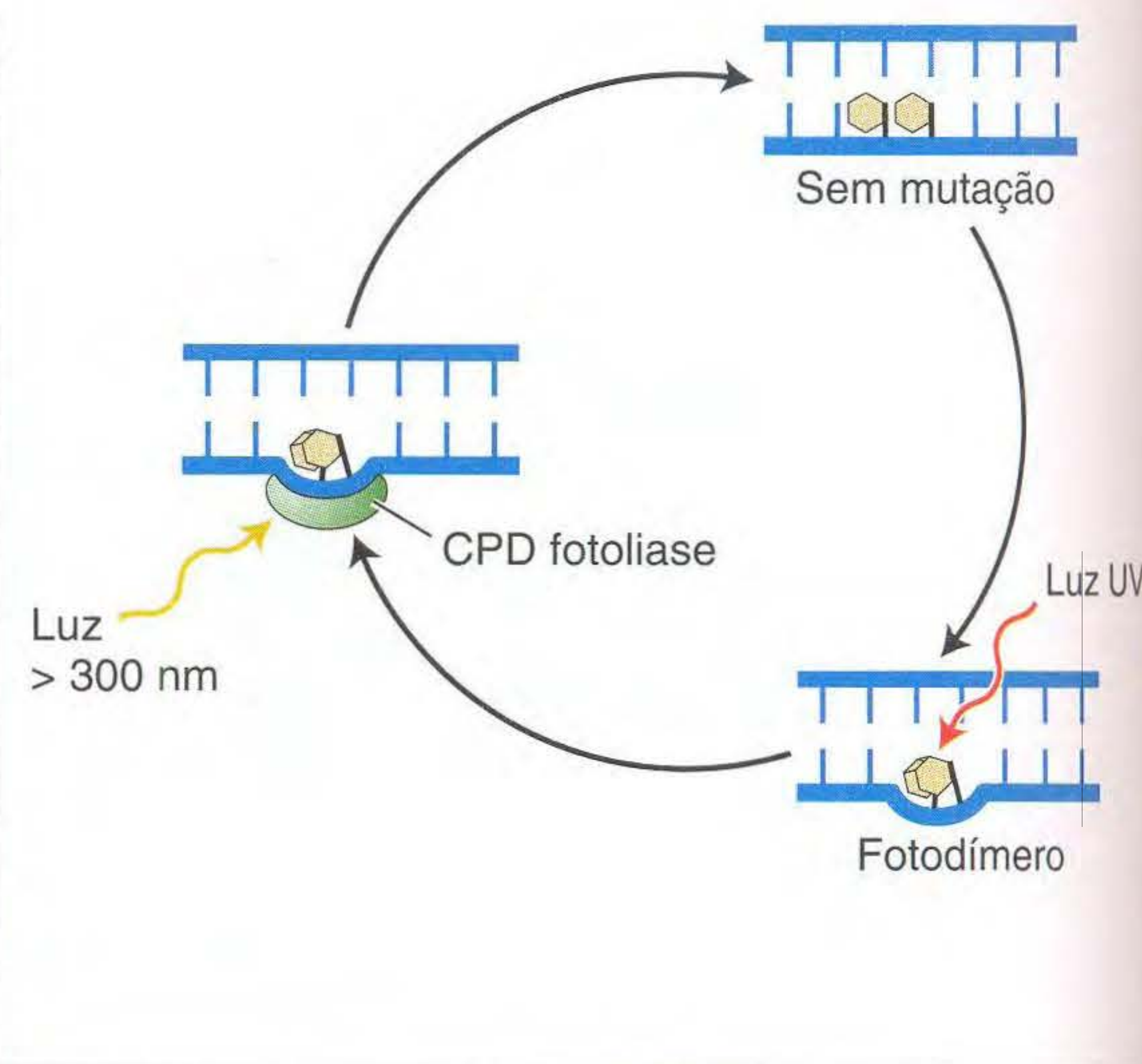


FIG. 15.22 A enzima CPD fotoliase divide o fotodímero de pirimidina ciclobutano para reparar essa mutação.

Essa enzima transfere o grupo metil de O-6-metilguanina para uma cisteína no sítio ativo da enzima. Entretanto, a transferência inativa a enzima, e, assim, esse sistema de reparo pode ser saturado se o nível de alquilação for suficientemente alto.

Reparo por excisão de base

Um importante princípio que orienta os sistemas genéticos celulares é o poder de complementariedade de uma sequência de nucleotídeos. (Lembre que a análise genética também depende muito desse princípio.) Importantes sistemas de reparo exploram as propriedades de complementariedade antiparalela para restaurar os segmentos danificados de DNA a seu estado inicial não-danificado. Nesses sistemas, uma base ou um segmento maior de cadeia de DNA é removido e substituído por um segmento de nucleotídeo recém-sintetizado complementar ao filamento-molde oposto.

Como esses sistemas dependem de complementariedade, ou homologia, do filamento-molde com o filamento que está sendo reparado, eles são chamados de **sistemas de reparo dependentes de homologia**. Ao contrário dos exemplos de reversão de danos descritos na seção anterior, essas vias incluem a remoção e a substituição de uma ou mais bases.

O primeiro sistema de reparo dependente de homologia que examinaremos é o sistema de **reparo por excisão de base**. Após a revisão do DNA pela DNA polimerase, o reparo por excisão de bases é o mecanismo mais importante usado para remover bases incorretas ou danificadas. O principal alvo do reparo por excisão de base é o dano não-volumoso a bases. Esse tipo de dano pode resultar de uma variedade de causas mencionadas nas seções anteriores, incluindo metilação, desaminação, oxidação ou perda espontânea de

Um pequeno dano de base é detectado e restaurado por reparo por excisão de base

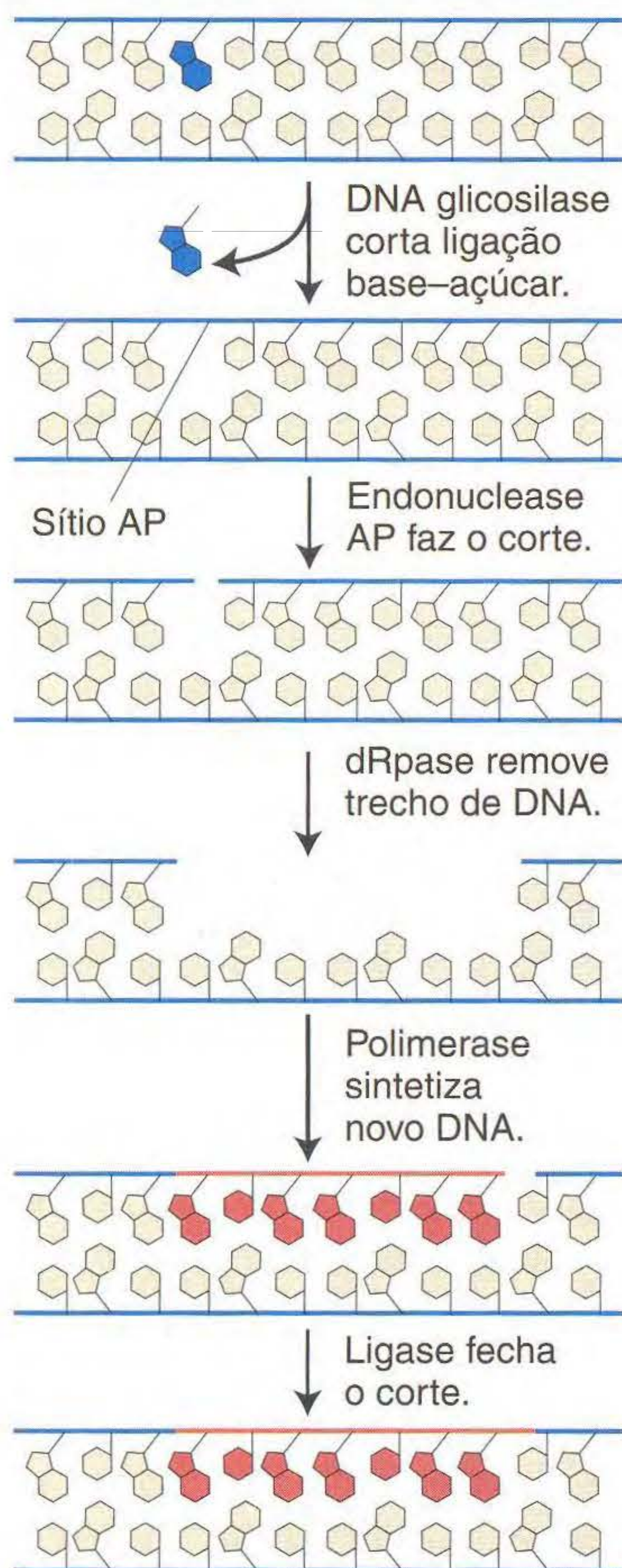


FIG. 15.23 No reparo por excisão de base, as bases danificadas são removidas e reparadas pela ação sequencial de uma DNA glicosilase, AP endonuclease, desoxirribosofosfodiesterase (dRpase), DNA polimerase e ligase.

uma base do DNA. O reparo por excisão de base (Figura 15.23) é feito pelas DNA glicosidasas que cortam ligações base-açúcar, liberando assim as bases alteradas e gerando sítios apurínicos ou apirimidínicos (AP). Uma enzima chamada AP endonuclease então corta o filamento danificado antes do sítio AP. Uma terceira enzima, desoxirribosofosfodiesterase, limpa o arcabouço removendo um trecho do açúcar-fosfato vizinho, de modo que a DNA polimerase pode preencher o espaço com nucleotídeos complementares ao outro filamento. A DNA ligase então une o novo nucleotídeo ao arcabouço (veja Figura 15.23).

Existem várias DNA glicosilases. Uma, uracil-DNA glicosilase, remove uracil do DNA. As unidades uracil que resultam de desaminação espontânea de citosina (veja anteriormente) podem levar a uma transição C para T caso não reparadas. Uma vantagem de ter timina (5-metiluracil) em vez de uracil como parceiro natural de pareamento de adenina no DNA é que os eventos de desaminação espontânea de citosina podem ser reconhecidos como anormais e, então, removidos e reparados. Se uracil fosse um constituinte normal do DNA, tal reparo não seria possível.

Entretanto, a desaminação cria outros problemas para bactérias e eucariontes. Analisando um grande número de mutações no gene *lacI*, Jeffrey Miller identificou locais no gene onde uma ou mais bases eram propensas a freqüentes mutações. Miller descobriu que os chamados pontos quentes mutacionais correspondem à desaminação de algumas citosinas. A análise da sequência de DNA da transição dos pontos quentes $G \cdot C \rightarrow A \cdot T$ no gene *lacI* mostrou que as 5-metilcitosinas estão presentes em cada ponto quente. A metilação do DNA eucariótico foi discutida no Capítulo 11. Similarmente, *E. coli* e outras bactérias também metilam seu DNA, embora para fins diferentes. Alguns dos dados desse estudo de *lacI* são mostrados na Figura 15.24. A altura de cada barra no gráfico representa a freqüência de mutações em cada sítio. As posições de 5-metilcitosinas pode ser bem correlacionada com a maioria dos sítios mutáveis.

5-Metilcitosina é um ponto quente para mutação

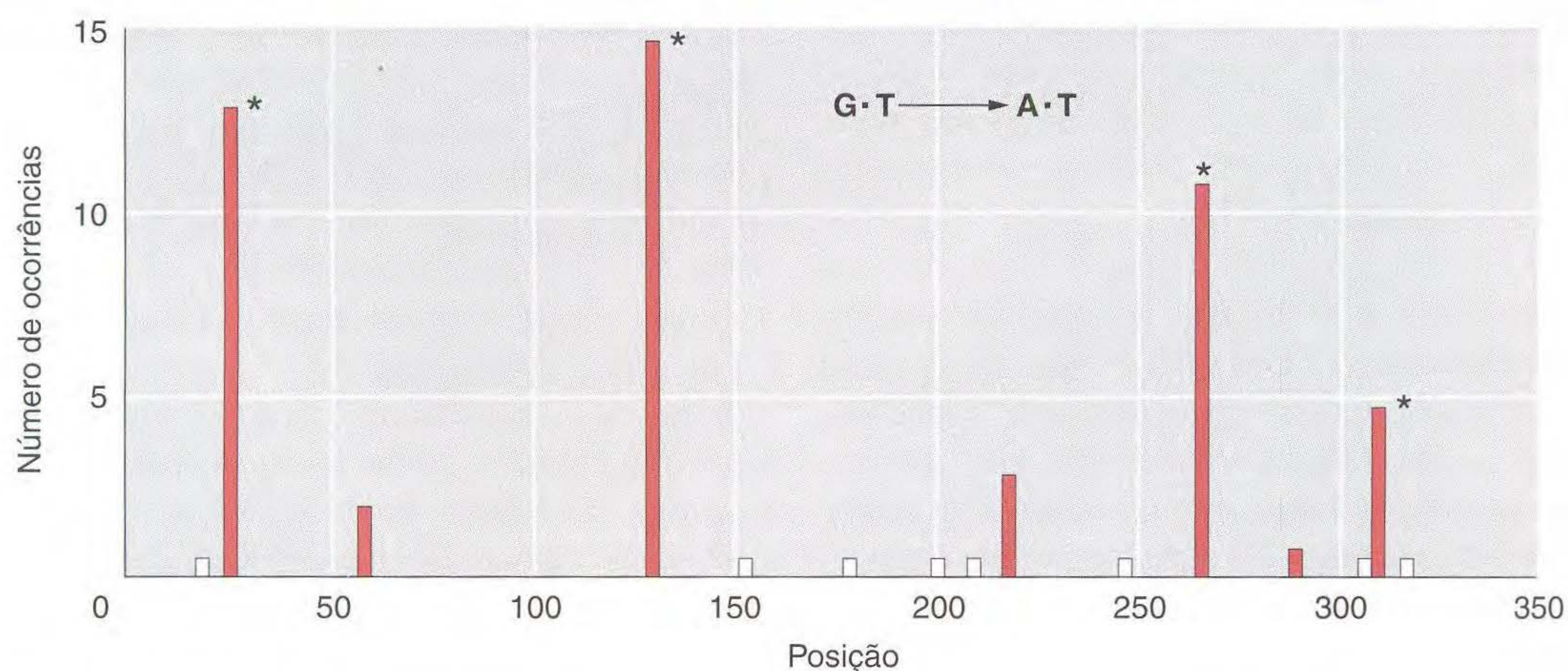
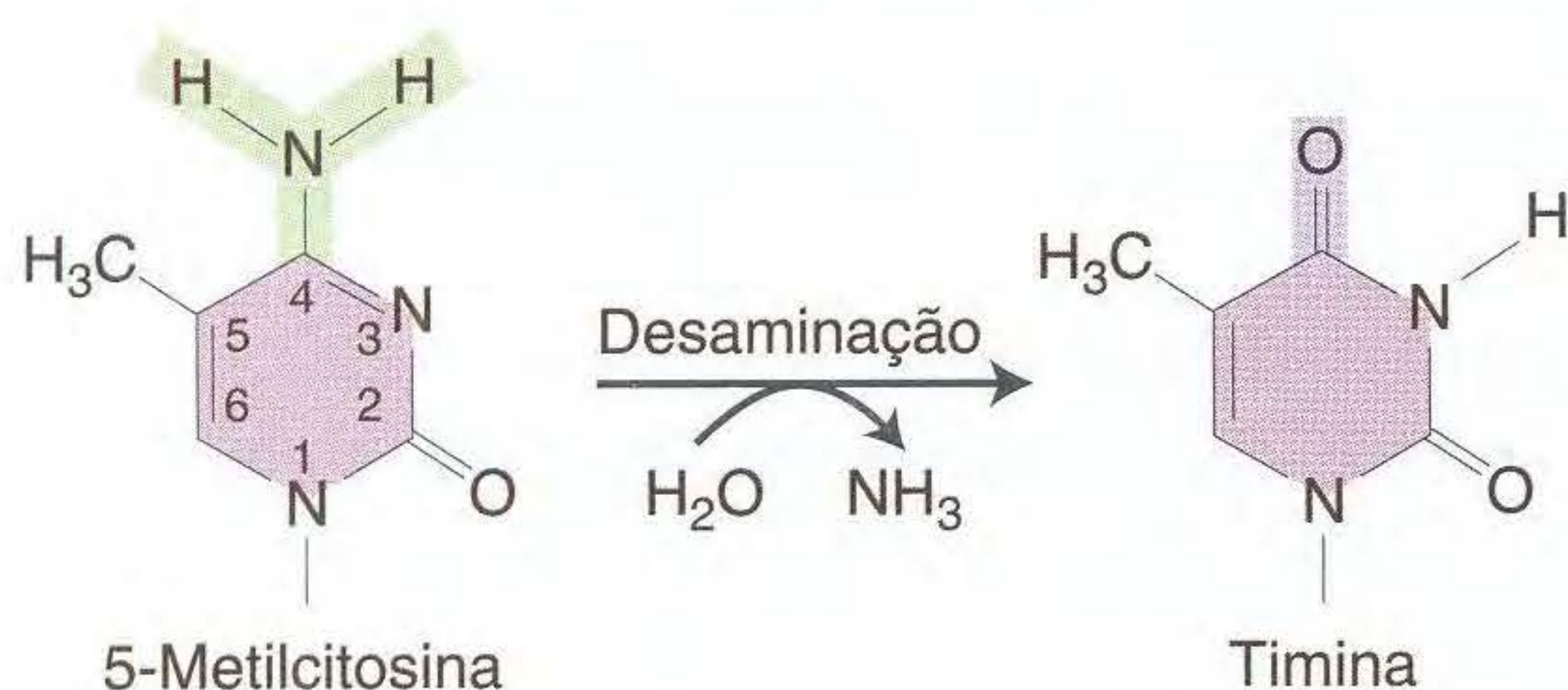


FIG. 15.24 Pontos quentes de metilcitosina em *E. coli*. Mutações sem sentido em 15 sítios diferentes em *lacI* foram registradas. Todas resultaram em transições $G \cdot C \rightarrow A \cdot T$. O asterisco (*) marca as posições de 5-metilcitosinas, e as barras brancas marcam os sítios onde estão as transições conhecidas como ocorrendo não-isoladas nesse grupo. [De C. Coulondre, J.H. Miller, P.J. Farabaugh, and W. Gilbert, "Molecular Basis of Base Substitution Hotspots in *Escherichia coli*," *Nature* 274, 1978, 775.]

Por que as 5-metilcitosinas são pontos quentes para mutações? A desaminação de 5-metilcitosina gera timina (5-metiluracil).



A timina não é reconhecida pela enzima uracil-DNA glicosilase e, portanto, não é reparada. Assim, as transições C → T geradas pela desaminação são vistas mais frequentemente nos sítios de 5-metilcitosina porque escapam a esse sistema de reparo. Uma consequência da freqüente mutação de 5-metilcitosina para timina é a sub-representação de dinucleotídeos CpG nas células eucarióticas, pois essa sequência é metilada para dar 5-metil-CpG, que é gradualmente, ao longo do tempo evolutivo, convertida em TpG.

No reparo por excisão de base, os danos não-volumosos ao DNA são reconhecidos por uma de várias enzimas chamadas DNA glicosilases que clivam as ligações base-açúcar, liberando a base incorreta. O reparo consiste na remoção do sítio, que agora não tem uma base, e na inserção da base correta orientada pela base complementar no filamento não-danificado.

Reparo por excisão de nucleotídeo

Embora a grande maioria dos danos sofridos por um organismo seja um pequeno dano que pode ser lidado pelo reparo de excisão de bases, esse mecanismo não pode corrigir adições volumosas que distorcem a hélice do DNA, adições tais como os dímeros de pirimidina de ciclobutano causadas por luz UV (veja Figura 15.18), nem corrigir danos a mais de uma base. Uma DNA polimerase não pode continuar a síntese de DNA por essas lesões, e, assim, o resultado é um bloqueio na replicação. Uma forquilha de replicação bloqueada pode causar a morte celular. Similarmente, uma base anormal ou danificada pode parar o complexo de transcrição. Para lidar com ambas essas situações, os procariontes e eucariontes usam uma via extremamente versátil chamada de **reparo por excisão de nucleotídeo (NER)**, que é capaz de desfazer os bloqueios de replicação e transcrição e reparar o dano.



Curiosamente, duas doenças autossômicas recessivas em humanos, **xeroderma pigmentoso (XP)** e **síndrome de Cockayne**, são causadas por defeitos no reparo de excisão de nucleotídeo. Embora os pacientes com XP ou síndrome de Cockayne sejam excepcionalmente sensíveis à luz UV, outros sintomas são marcadamente diferentes. O xeroderma pigmentoso foi introduzido no início deste capítulo, e é caracterizado pelo desenvolvimento precoce de cânceres, especialmente câncer de pele e, em alguns casos,

defeitos neurológicos. Em contraste, os pacientes afetados pela síndrome de Cockayne têm uma variedade de distúrbios desenvolvimentais, incluindo nanismo, surdez e retardo. Em termos amplos, os pacientes com XP têm um câncer precoce, enquanto os pacientes com síndrome de Cockayne envelhecem prematuramente. Como podem os defeitos na mesma via de reparo levar a sintomas de doenças tão diferentes? Embora não exista uma resposta simples para essa pergunta, os trabalhos na base genética dessas doenças levaram à identificação de proteínas importantes na via NER.

O reparo por excisão de nucleotídeo é um processo complexo que requer dúzias de proteínas. Apesar dessa complexidade, o processo de reparo pode ser dividido em quatro fases:

1. Reconhecimento das bases danificadas
2. Montagem de um complexo multiproteico no local
3. Cortar o filamento danificado em local antecedente e posterior ao sítio danificado e remoção dos nucleotídeos (~ 30) entre os cortes.
4. Uso do filamento não-danificado como molde para a DNA polimerase seguido de ligação do filamento.

O fato de, como já foi mencionado, a parada da forquilha de replicação e a parada dos complexos de transcrição ativarem essa via de reparo significa que existem dois tipos de reparo por excisão de nucleotídeos que diferem no reconhecimento do dano (etapa 1). Sabemos, agora, que o tipo que repara regiões transcritas de DNA é chamado, não-surpreendentemente, de **reparo por excisão de nucleotídeo acoplado à transcrição (TC-NER)**. O outro tipo, o **reparo genômico global (GGR)**, corrige lesões em qualquer parte do genoma, e é ativado por forquilhas de replicação paradas. Como podemos ver na Figura 15.25, embora a etapa de reconhecimento difira, tanto TC-NER quanto GGR compartilham as etapas de 2 a 4.

Neste ponto, alguns de vocês devem estar indagando: Que diferenças nos sintomas de doença de XP e síndrome de Cockayne eram devidas a mutações em classes diferentes de proteínas de reconhecimento? Vocês estariam no caminho certo em fazerem essa pergunta. Os pacientes com a síndrome de Cockayne têm uma mutação em uma de duas proteínas chamadas CSA e CSB, que são consideradas para reconhecer complexos de transcrição parados. Um modelo de como CSA e CSB podem ligar-se à polimerase travada e, então, atrair o complexo multiproteico necessário para o reparo do DNA é mostrado na Figura 15.25.

Para GGR, a primeira etapa é o reconhecimento de uma base que (1) foi danificada por um complexo heterodímero e, assim, distorce a dupla hélice e (2) não é parte de um molde ativamente transcrito. Esse complexo de reconhecimento atrai o complexo multiproteico, que contém as 10 subunidades do fator geral de transcrição TFIIH. Duas das 10 subunidades, XPB e XPD, são helicases que desenrolam o DNA, e toda a estrutura é estabilizada por RPA, uma proteína de ligação ao DNA unifilamentar. As etapas subsequentes que medeiam a clivagem e excisão da base danificada e até

Duas vias de reparo por excisão de nucleotídeo

Reparo genômico global (GGR)

NER acoplado a transcrição

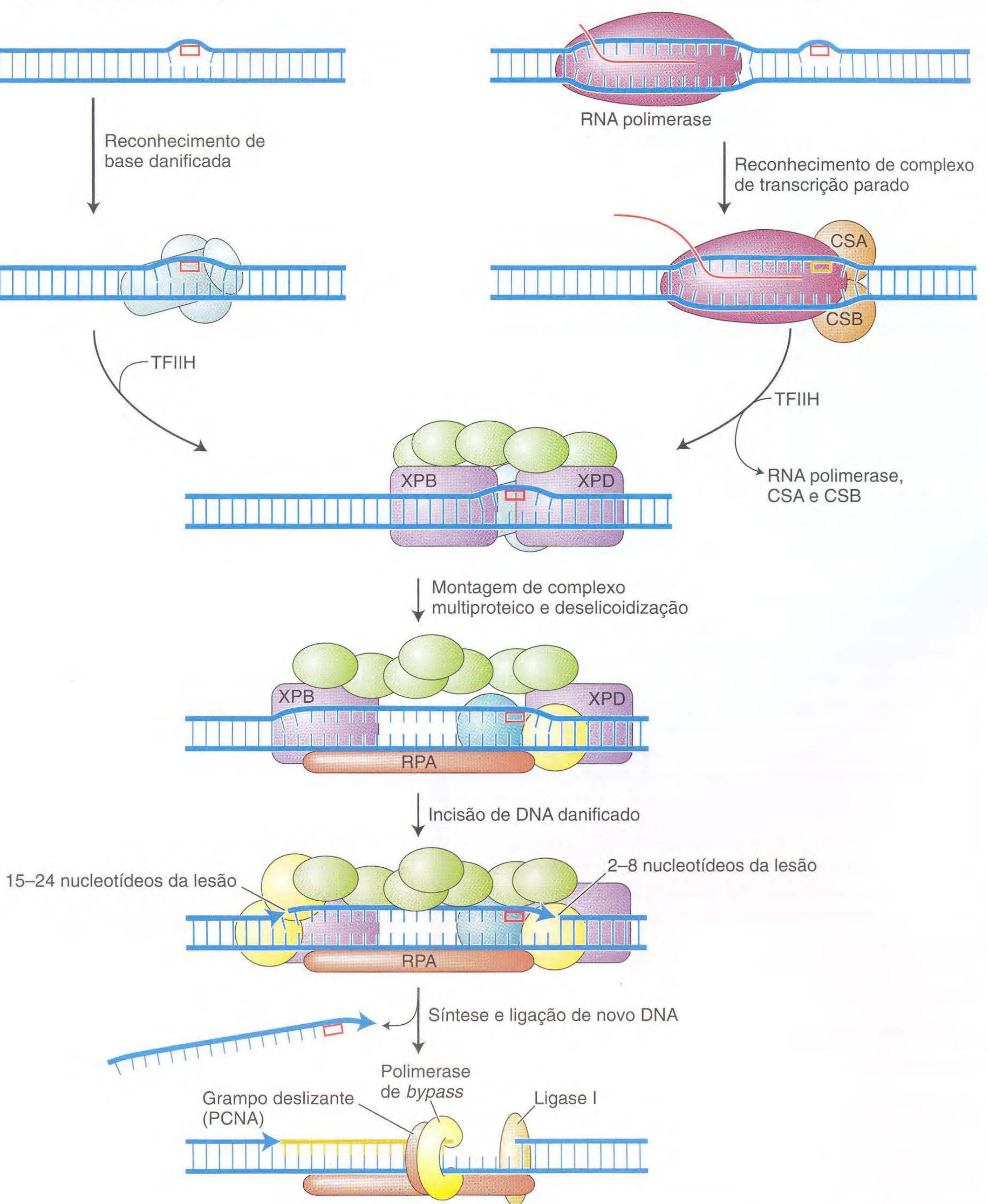


FIG. 15.25 A via de reparo por excisão de nucleotídeo é ativada quando adições volumosas ou múltiplas bases danificadas são reconhecidas em regiões não-transcritas (GGR) ou transcritas (TC-NER) do genoma. Um complexo multiproteico corta várias bases e as produz novamente usando o filamento oposto como molde. Veja o texto para detalhes.

30 nucleotídeos adjacentes seguidos da síntese de DNA para preencher o espaço são comuns às vias TC-NER e GGR. As etapas que resintetizam o DNA para preencher o espaço (chamada síntese de reparo de espaço) serão discutidas em

mais detalhe mais adiante, pois outras vias de reparo compartilham essa etapa.

Muitas das contrapartes humanas das proteínas nucleares de reparo por excisão (XPA, XPB, e assim por diante)

foram primeiro identificadas em pacientes com XP. Tanto XP quanto a síndrome de Cockayne são exemplos de **distúrbios genéticos heterogêneos** nos quais os pacientes têm mutações em um dentre vários genes com papéis em um processo particular — nesse caso, reparo por excisão de nucleotídeos.

As diferenças moleculares entre TC-NER e GGR podem dar uma explicação para sintomas diferentes apresentados por pacientes com XP e síndrome de Cockayne? Lembre que os pacientes com XP desenvolvem cânceres precoces, enquanto os pacientes da síndrome de Cockayne têm uma variedade de sintomas associados ao envelhecimento prematuro. Vimos que o sistema de reparo dos pacientes com síndrome de Cockayne não pode reconhecer complexos de transcrição travados. Como consequência desse defeito, a célula é mais propensa a ativar a via de suicídio de apoptose. Em uma pessoa saudável, a morte celular é geralmente preferível à propagação de uma célula que tem um dano ao DNA mantido. Entretanto, de acordo com essa teoria, a via de morte celular seria ativada mais frequentemente em um paciente com síndrome de Cockayne, levando assim a uma variedade de sintomas de envelhecimento prematuro. Em contraste, os pacientes com XP podem reconhecer complexos de transcrição travados (eles têm proteínas normais CSA e CSB) e evitam a morte celular quando a transcrição é reiniciada. Entretanto, eles não podem reparar o dano original devido a mutações em uma de suas proteínas XP. Assim, uma das mutações se acumularia nas células de pacientes com XP e, como já foi dito neste capítulo, a presença de mutações, quer causadas por mutágenos, quer pela falha nas vias de reparo, aumenta o risco de desenvolvimento de muitos tipos de câncer.

O reparo por excisão de nucleotídeo é uma via versátil que reconhece e corrige as lesões do DNA devidas amplamente a danos de UV, e, ao fazer isto, alivia forquilhas de replicação travadas e complexos de transcrição. Os pacientes com xeroderma pigmentoso e síndrome de Cockayne são sensíveis a UV devido a mutações em proteínas-chave de reparo de excisão de nucleotídeo que reconhecem ou reparam bases danificadas.

Reparo pós-replicação: reparo de mal pareamento

Você aprendeu, na primeira metade deste capítulo, que muitos erros ocorrem na replicação do DNA. De fato, a taxa de erro é de cerca de 10^{-5} . A correção pela função de revisão 3' para 5' da polimerase replicativa reduz a taxa de erro a menos de 10^{-7} . A principal via que corrige os erros replicativos restantes é chamada de **reparo de mal pareamento**. Essa via de reparo reduz a taxa de erro a menos de 10^{-9} , reconhecendo e reparando bases mal pareadas e pequenas alças causadas pela inserção e deleção de nucleotídeos (indels) no curso da replicação. Por esses valores, você pode ver que as mutações que levam à perda do reparo de mal pareamento podem aumentar a frequência de mutação em 100 vezes. De fato, a perda do reparo de mal pareamento está associada a formas hereditárias de câncer de cólon.

Os sistemas de reparo de mal pareamento têm que fazer pelo menos três coisas:

1. Reconhecer bases mal pareadas
2. Determinar que base está mal pareada
3. Remover a base incorreta e fazer a síntese de reparo

A maioria do que se sabe sobre o reparo de mal pareamento vem de décadas de análise genética e bioquímica com o uso da bactéria modelo *E. coli* (veja o boxe Organismo-modelo *E. coli* no Capítulo 5). Especialmente notável foi a reconstituição do sistema de reparo no tubo de ensaio no laboratório de Paul Modrich. A conservação de muitas das proteínas de reparo de mal pareamento, desde bactérias e leveduras até humanos, indica que essa via é tanto antiga quanto importante em todos os organismos vivos. Recentemente, o sistema de reparo de mal pareamento humano também foi reconstituído no tubo de ensaio no laboratório de Modrich. A habilidade em estudar os detalhes da reação estimulará futuros estudos da via humana. Entretanto, por enquanto enfocaremos o bem caracterizado sistema de *E. coli* (Figura 15.26).

A primeira etapa no sistema de mal pareamento é o reconhecimento do dano no DNA recém-replicado pela proteína MutS. A ligação dessa proteína a distorções na dupla hélice de DNA causada por bases pareadas erradas inicia a via de reparo de mal pareamento atraindo três outras proteínas para o sítio da lesão (MutL, MutH e UvrD). A proteína principal é MutH, que desempenha a função crucial de cortar o filamento contendo a base incorreta. Sem essa habilidade para discriminar entre as bases corretas e incorretas, o sistema de reparo de mal pareamento não pode determinar qual base remover para evitar que surja uma mutação. Se, por exemplo, um reparo de mal pareamento G-T ocorre como um erro de replicação, como o sistema pode determinar se G ou T é a incorreta? Ambas são bases normais no DNA. Mas os erros de replicação produzem pareamentos errados no filamento recém-sintetizado, e, assim, o sistema de mal pareamento substitui a base nesse filamento.

Como o reparo de mal pareamento distingue o filamento recém-sintetizado do antigo? Lembre, do Capítulo 11, que as bases citosina em geral são metiladas nos eucariontes, e que essa marca epigenética é propagada do filamento parental para o filho logo após a replicação. O DNA de *E. coli* também é metilado, mas os grupos metil relevantes para o reparo de mal pareamento são adicionados a bases adenina. Para distinguir o filamento-molde antigo do filamento recém-sintetizado, o sistema de reparo bacteriano tira proveito de um retardo na metilação da seguinte sequência:

5'-G-A-T-C-3'

3'-C-T-A-G-5'

A enzima metilante é a adenina metilase, que cria 6-metiladenina em cada filamento. Entretanto, a adenina metilase requer vários minutos para reconhecer e modificar trechos GATC recém-sintetizados. Nesse intervalo, a proteína MutH corta o sítio de metilação no filamento contendo a A que ainda não foi metilada. Esse sítio pode

O reparo de mal pareamento corrige erros replicativos

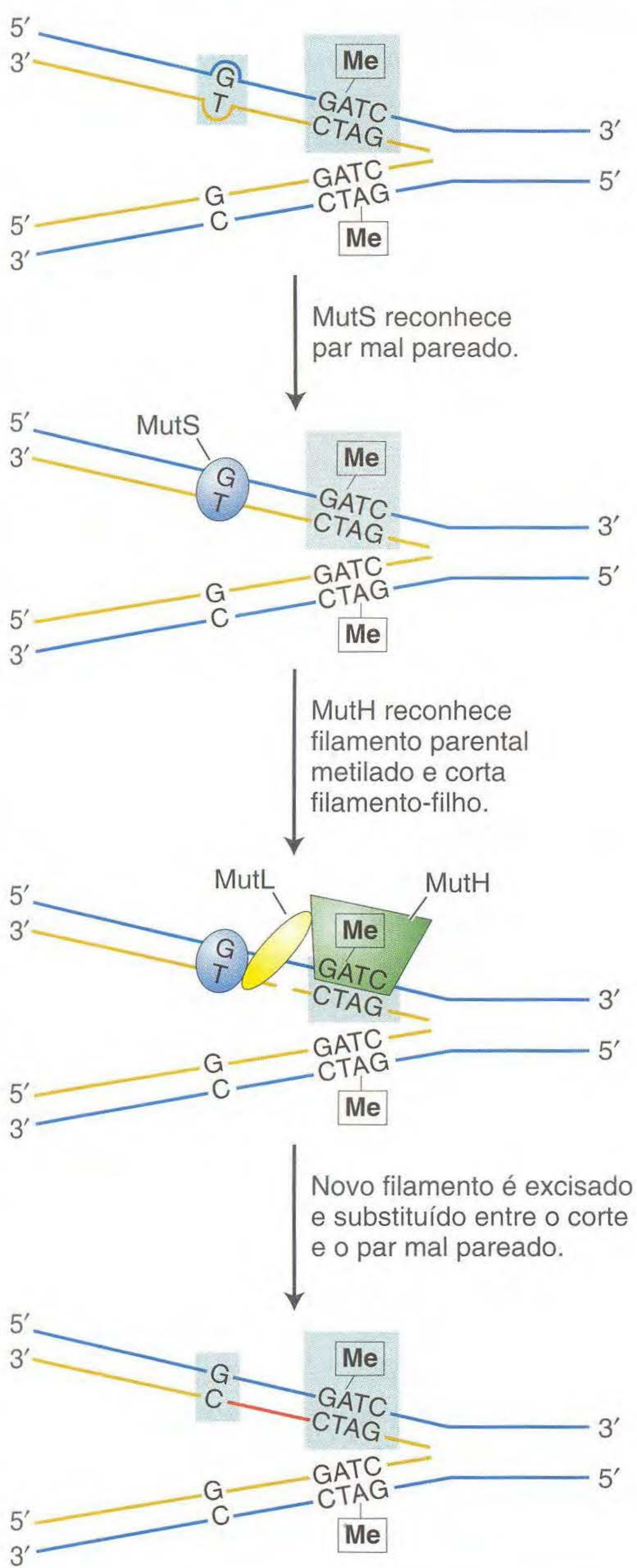


FIG. 15.26 Modelo para reparo de mal pareamento em *E. coli*. O DNA é metilado (Me) em uma A em uma sequência GATC. A replicação do DNA produz um dúplice hemimetilado que existe até que a metilase possa modificar o filamento recém-sintetizado. O sistema de reparo de mal pareamento faz quaisquer correções necessárias com base na sequência encontrada no filamento metilado (molde original). MutS, MutH e MutL são proteínas.

distar centenas de pares de bases da base erradamente pareada. Após o sítio ter sido cortado, a proteína UrvD liga-se ao corte e usa sua atividade de helicase para desenrolar o DNA. Uma proteína de ligação unifilamentar protetora reveste e desenrola o filamento parental, enquanto a parte do novo filamento entre o pareamento errado e o corte é removida.

Muitas das proteínas no reparo de mal pareamento em *E. coli* são conservadas no reparo em humanos. Entretanto, não se conhece como os eucariontes reconhecem e reparam apenas o filamento recém-replicado. O problema é particularmente perplexo em organismos em que há ausência da maioria ou de toda a metilação do DNA, tais como leveduras, *Drosophila* e *C. elegans* (veja o Capítulo 11). Um modelo popular propõe que a discriminação é baseada no reconhecimento das pontas livres 3' que caracterizam os filamentos de replicação contínua e descontínua recém-sintetizados.

Um alvo importante do sistema de mal pareamento humano são as curtas sequências repetidas que podem ser expandidas ou deletadas na replicação pelo mecanismo de mal pareamento descrito previamente (veja Figura 15.10). As mutações em alguns componentes dessa via foram mostrados como sendo responsáveis por várias doenças humanas, especialmente cânceres. Existem milhares de repetições curtas (microssatélites) situadas ao longo do genoma humano (veja o Capítulo 4). Quase todas estão localizadas em regiões não-codificantes (considerando que a maioria do genoma não é codificante), algumas situadas em genes que são críticos para o crescimento normal e o desenvolvimento.



Portanto, os defeitos na via de reparo de mal pareamento de humanos seriam previstos tendo várias doenças graves como consequência. Essa previsão mostrou-se verdadeira, e um caso é a síndrome chamada de câncer colorretal não-polipose hereditária (HNPCC). A despeito de seu nome, não é propriamente um câncer, mas aumenta o risco de câncer. Uma das mais comuns predisposições herdadas ao câncer, a doença afeta até 1 em cada 200 pessoas no mundo ocidental. Os estudos mostraram que HNPCC resulta da perda do sistema de reparo de mal pareamento devido, em grande parte, a mutações herdadas em genes que codificam as contrapartes humanas (e homólogos) das proteínas bacterianas MutS e MutL (veja Figura 15.26). A herança de HNPCC é autossômica dominante. As células com uma cópia funcional dos genes de reparo de mal pareamento têm atividade de reparo normal de mal pareamento, mas as linhagens de células tumorais surgem de células que perderam a cópia funcional e são, portanto, deficientes do mal pareamento. Essas células têm altas taxas de mutação devido, em parte, à incapacidade em corrigir a formação de indels na replicação.

O sistema de reparo de mal pareamento corrige erros na replicação que não são corrigidos na função de revisão da DNA polimerase replicativa. O reparo é restrito ao filamento recém-sintetizado, que é reconhecido pela maquinaria de reparo em procariontes porque não tem um marcador de metilação.

Reparo propenso a erro: síntese de DNA translesão

Até aqui, todos os mecanismos de reparo que encontramos são livres de erro, pois eles ou reverterem diretamente o dano, ou usam a complementariedade de bases para inserir a base

correta. Existem vias de reparo que são uma fonte significativa de mutação. Esses mecanismos parecem ter evoluído para evitar a ocorrência de resultados potencialmente mais sérios, tais como morte celular ou câncer. Como já foi mencionado, uma forquilha de replicação parada pode iniciar a via de morte celular. Tanto em procariontes quanto em eucariontes, tais bloqueios de replicação podem ser *contornados* pela inserção de bases inespecíficas. Em *E. coli*, esse processo requer a ativação do **sistema SOS**. O nome SOS vem da idéia de que esse sistema é induzido como uma resposta de emergência para evitar a morte celular na presença de dano significativo ao DNA. Como tal, a indução SOS é um mecanismo de último recurso, uma forma de tolerância ao dano que permite que a célula troque a morte por um certo nível de mutagenese.

Levou mais de 30 anos para descobrir como o sistema SOS gera mutações enquanto permite que a DNA polimerase con-

torne a lesão na forquilha de replicação parada. Já estamos familiarizados com os danos ao DNA induzidos pela luz UV (veja Figura 15.18). Uma classe incomum de mutantes de *E. coli* que sobreviveu à exposição a UV sem manter mutações adicionais foi isolada nos anos de 1970. O fato de que tais mutantes existiram sugeriu que alguns genes de *E. coli* funcionam para gerar mutações quando expostos à luz UV. As mutações induzidas por UV não ocorrerão se os genes *DinB*, *UmuC* e *UmuD'* estiverem mutados.

A Figura 15.27 mostra as etapas no mecanismo SOS. Na primeira etapa, a luz UV induz a síntese de uma proteína chamada RecA. Voltaremos a falar sobre a proteína RecA mais adiante no capítulo porque ela é um componente-chave nos mecanismos de reparo de DNA e recombinação. Quando a polimerase replicativa (DNA polimerase III) pára em um sítio de DNA danificado, o DNA adiante da polimerase continua a ser desenrolado, expondo regiões de DNA unifilamentar que se tornam ligados por proteínas de ligação a filamentos únicos. Em seguida, as proteínas RecA juntam-se a proteínas de ligação unifilamentar e formam um filamento proteína-DNA. O filamento RecA é a forma biologicamente ativa desta proteína. Nessa situação, RecA atua como um sinal que leva à indução de vários genes hoje conhecidos como codificando membros de uma família recém-descoberta de DNA polimerases que podem contornar o bloqueio de replicação e são distintas das polimerases replicativas. As DNA polimerases que podem contornar as paradas de replicação também foram encontradas em diversos grupos de eucariontes, variando de leveduras a humanos. Essas polimerases eucarióticas contribuem para um mecanismo de tolerância a danos, chamado de **síntese de DNA translesão**, que se assemelha ao sistema SOS de *E. coli*.

Essas **polimerases translesão**, ou **bypass**, como são conhecidas, diferem das principais polimerases replicativas de vários modos. Primeiro, elas podem tolerar adições incomumente grandes de bases. Enquanto a polimerase replicativa pára se uma base não se ajusta a um sítio ativo, as polimerases de *bypass* têm bolsos muito maiores que podem acomodar bases danificadas. Segundo, em algumas situações, as polimerases de *bypass* têm uma taxa de erro muito maior, em parte porque não têm atividade de revisão 3' para 5' das principais polimerases replicativas. Terceiro, elas podem adicionar apenas alguns nucleotídeos antes de sair. Essa característica é atraente porque a principal função de uma polimerase propensa a erro é desbloquear a forquilha de replicação, não a síntese de longos trechos de DNA que podem conter muitos pareamentos errados.

Várias polimerases de *bypass* que parecem estar sempre presentes em células eucarióticas são hoje conhecidas. Como elas estão sempre presentes, seu acesso ao DNA deve ser regulado, de modo a serem usadas apenas quando necessário. A célula desenvolveu uma eficiente solução para esse problema. Lembre que uma parte integral do replisomo é a proteína PCNA (antígeno nuclear de proliferação celular) que funciona como uma presilha deslizante para orquestrar a miríade de eventos na forquilha de replicação (veja Figura 7.18). Uma proteína crítica presente em

A síntese translesão ultrapassa as lesões nas forquilhas de replicação paradas

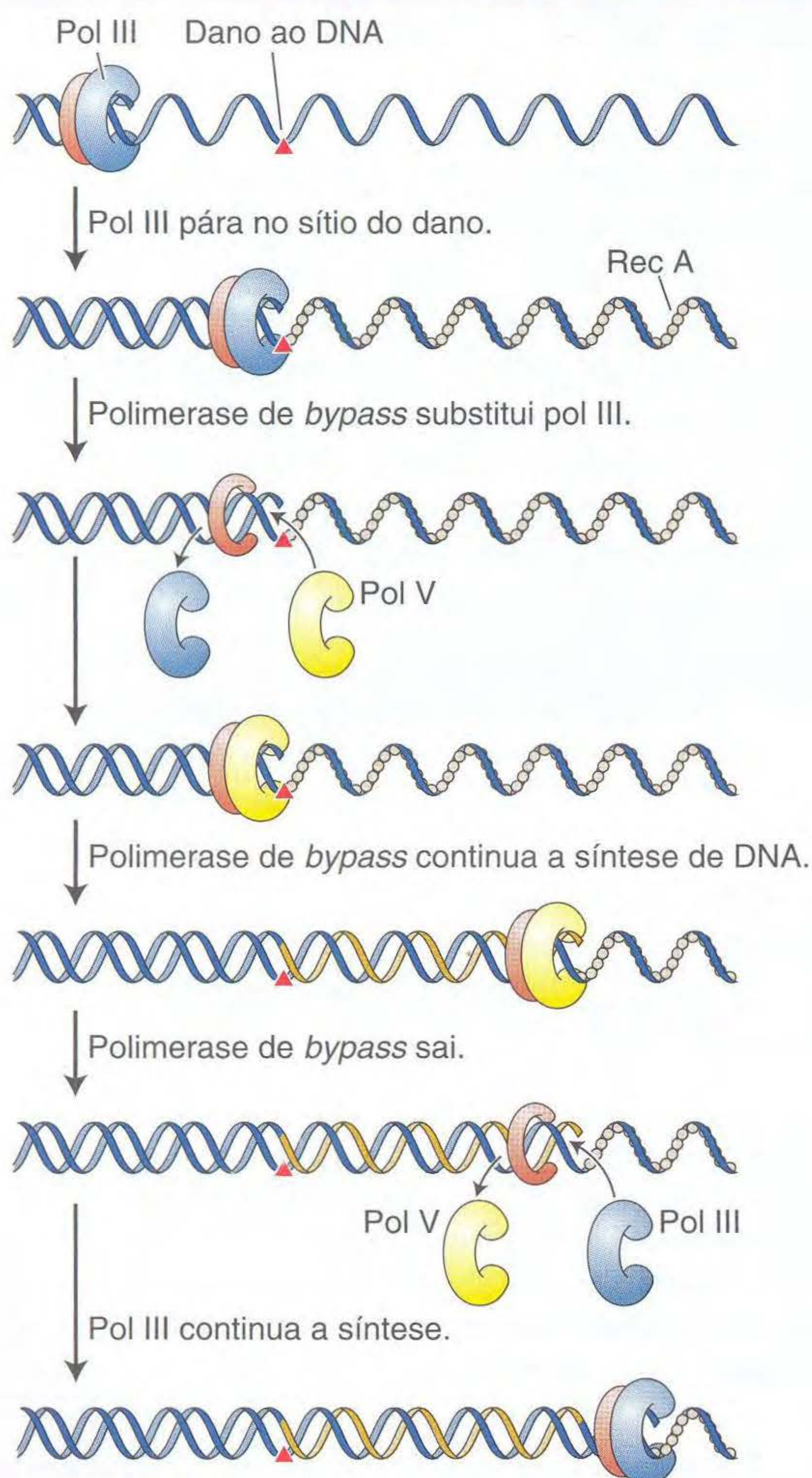


FIG. 15.27 Um modelo de síntese translesão em *E. coli*. No curso da replicação, a DNA polimerase III é temporariamente substituída por uma polimerase translesional (pol V) que pode continuar a replicar passando pela lesão. As polimerases translesionais são propensas a erro. [De E.C. Friedberg, A.R. Lehmann, and R.P. Fuchs, "Trading Places: How Do DNA Polymerases Switch During Translesion DNA Synthesis?" *Molec. Cell* 18, 2005, 499–505.]

A mudança de polimerase requer a adição de um monômero de ubiquitina

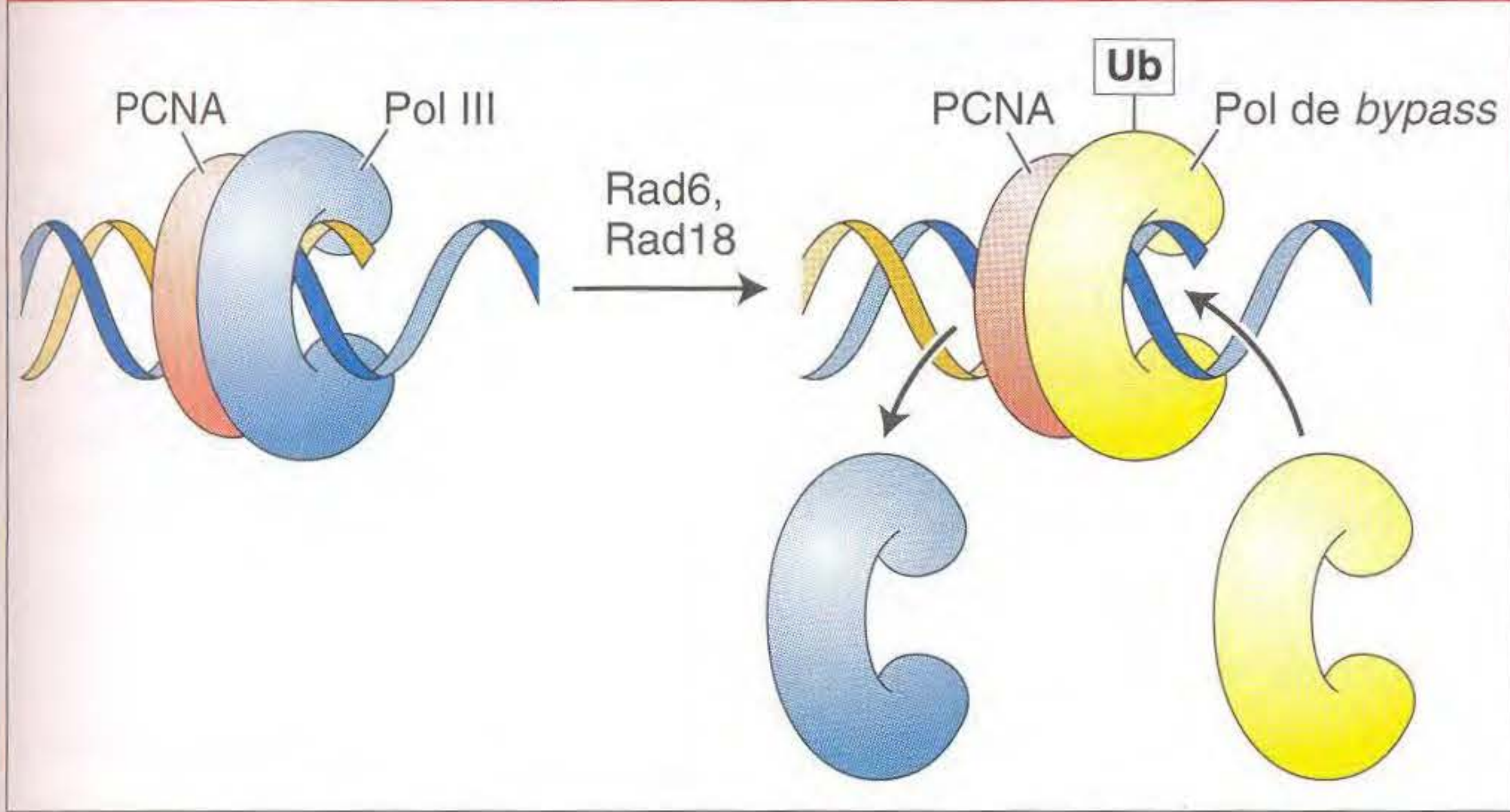


FIG. 15.28 A adição de um único monômero de ubiquitina (Ub) ao grupo deslizante (PCNA) permite que a polimerase translesão se ligue à PCNA e comece a replicação.

uma forquilha de replicação parada é Rad6, curiosamente uma enzima que adiciona ubiquitina às proteínas (Figura 15.28). Como descrito no Capítulo 9, a adição de cadeias a muitos monômeros de ubiquitina serve para marcar a proteína para degradação (veja Figura 9.23). Em contraste, a ligação de um único monômero de ubiquitina à PCNA muda sua conformação, de modo que ela agora pode ultrapassar a polimerase e orquestrar a síntese translesão. A remoção enzimática da marcação em PCNA leva à dissociação da polimerase de *bypass* e à eventual restauração da replicação normal. Qualquer mal pareamento de bases devido à síntese translesão ainda tem a chance de detecção e correção pela via de reparo de mal pareamento.

A regulação da função de PCNA pela adição e remoção dos monômeros de ubiquitina ilustra a importância das modificações pós-traducionais em eucariontes. Se o dano da base no filamento-molde não é rapidamente corrigido, a forquilha de replicação travada sinalizará a ativação da via de morte celular. Uma célula eucariótica não pode esperar a síntese *de novo* das polimerases de *bypass* após a transcrição e tradução como ocorre no sistema SOS de *E. coli*. Ao contrário, as polimerases de *bypass* eucarióticas são constitutivamente transcritas e estão sempre presentes. Seu acesso à forquilha de replicação é controlado por modificações pós-traducionais rápidas e reversíveis.

Na síntese translesão, as polimerases de *bypass* são recrutadas para a forquilha de replicação que parou devido ao dano no filamento-molde. As polimerases de *bypass* podem introduzir erros no curso da síntese que podem persistir e levar a uma mutação que pode ser corrigida por outros mecanismos, tais como o reparo de mal pareamento.

Reparo de quebras bifilamentares

Como já vimos, muitos sistemas de correção exploram a complementariedade do DNA para fazer reparos livres de erros. Tais reparos livres de erros são caracterizados por dois estágios: (1) remoção de bases danificadas, talvez juntamente

com DNA vizinho, de um filamento da dupla hélice e (2) uso do outro filamento como molde para a síntese de DNA necessária para preencher o espaço unifilamentar. Entretanto, o que ocorreria se *ambos* os filamentos da dupla hélice fossem danificados de tal modo que a complementariedade não pudesse ser explorada? Por exemplo, a exposição a raios X geralmente faz com que ambos os filamentos da dupla hélice sejam quebrados em sítios bem próximos. Esse tipo de mutação é chamado de **quebra bifilamentar**. Se não for reparada, a quebra bifilamentar pode causar uma variedade de anomalias cromossômicas que resultam na morte celular ou um estado pré-canceroso.

Curiosamente, a geração de quebras bifilamentares é uma característica integrante de alguns processos celulares normais que requerem rearranjos de DNA. Um exemplo é a recombinação meiótica. Como será visto no restante deste capítulo, a célula usa muitas das mesmas proteínas e vias para reparar quebras bifilamentares e efetuar a recombinação meiótica. Por esse motivo, começamos enfocando os mecanismos moleculares que reparam quebras bifilamentares antes de voltar nossa atenção para o mecanismo de recombinação meiótica.

As quebras bifilamentares podem surgir espontaneamente (por exemplo, em resposta a espécies de oxigênio reativo produzido como subproduto do metabolismo celular) ou podem ser induzidas por radiação ionizante. São conhecidos vários mecanismos para reparar quebras bifilamentares, e novos mecanismos ainda estão sendo descobertos. Dois mecanismos distintos são descritos na seção seguinte: junção de pontas não-homólogas e recombinação homóloga.

Junção de pontas não-homólogas Muitos dos mecanismos de reparo previamente descritos ocorrem na fase S do ciclo celular, quando o DNA está se replicando na preparação para a mitose ou meiose. Entretanto, ao contrário das células da maioria dos procariontes e eucariontes inferiores, as células dos eucariontes superiores geralmente não estão replicando seu DNA, pois estão na fase de repouso do ciclo celular ou pararam totalmente de se dividir. O que ocorre quando a quebra bifilamentar ocorre em células onde os filamentos não-danificados ou cromátides-irmãs não estão presentes? A resposta é que essas pontas devem ser reparadas, seja perfeita ou imperfeitamente, porque as pontas quebradas podem iniciar rearranjos cromossômicos potencialmente perigosos que podem levar a um estado canceroso (veja o Capítulo 16).

Um dos modos pelos quais os eucariontes colocam as pontas bifilamentares quebradas juntas novamente é pelo mecanismo não-elegante mas importante chamado de **junção de pontas não-homólogas (NHEJ)**, que é mostrado na Figura 15.29. Como em outros mecanismos de reparo, a primeira etapa na via NHEJ é o reconhecimento do dano. A via NHEJ é iniciada quando duas proteínas muito abundantes, KU70 e KU80, ligam-se às pontas quebradas, formando um heterodímero que serve a duas funções: primeira, impede outros danos às pontas e, segunda, recruta outras proteínas que aparam as pontas do filamento para gerar as pontas 5'-P

A junção de pontas não-homólogas propensa a erro repara quebras bifilamentares

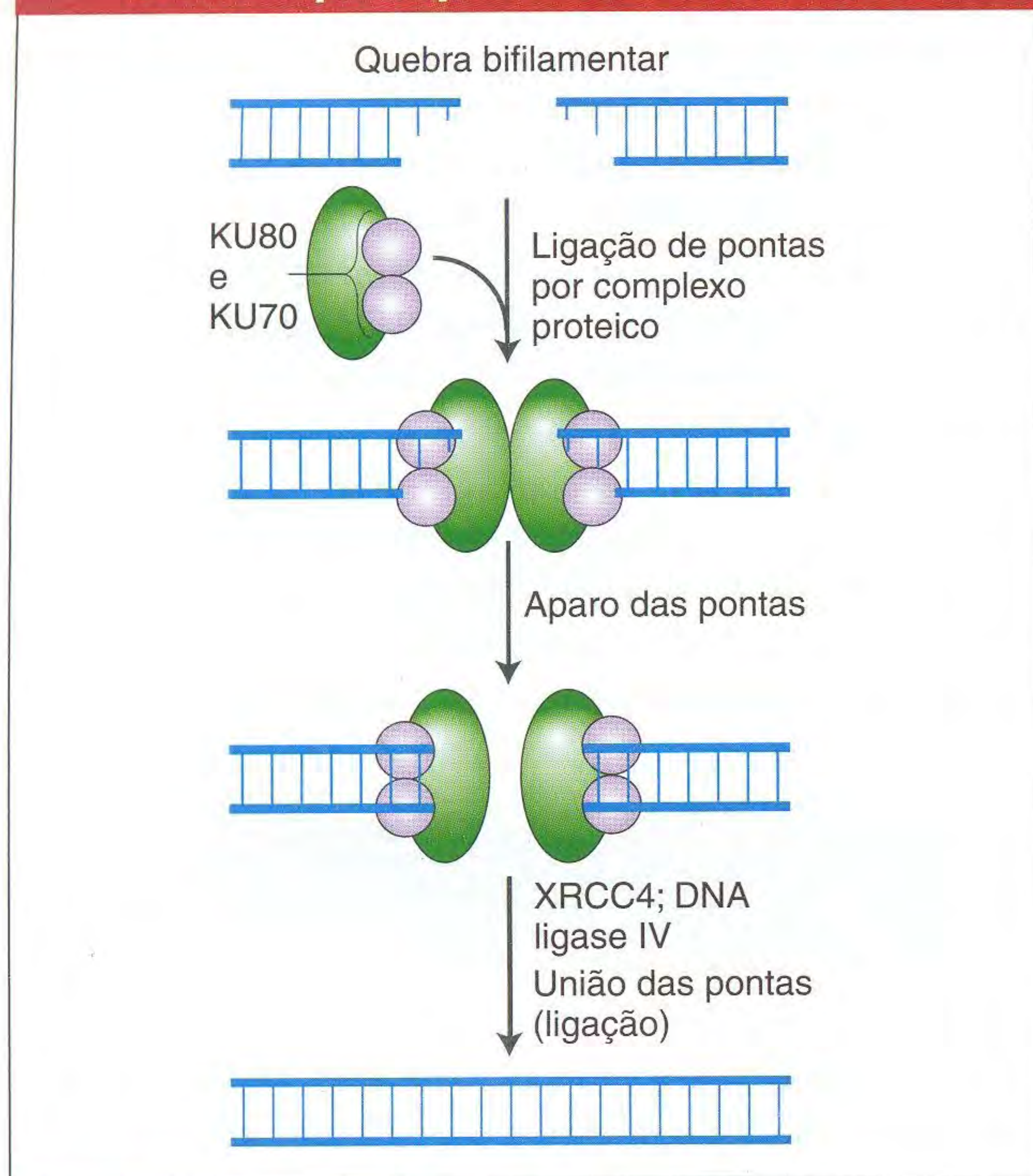


FIG. 15.29 Mecanismo de junção de pontas não-homólogas (NHEJ). Esse mecanismo é propenso a erro. Veja o texto para mais detalhes.

e 3'-OH que são necessárias para a ligação. A DNA ligase IV então junta as duas pontas.

Como os cientistas sabem quando todos os componentes de uma via biológica foram identificados? Sabemos que esse problema é difícil. A identificação de um componente novo da via NHEJ forneceu um exemplo recente.

Por vários motivos, todos os componentes da via NHEJ foram considerados como tendo sido identificados. Entretanto, os geneticistas analisando uma linhagem celular (chamada de 2BN) derivada de uma criança com um distúrbio raro herdado tiveram uma surpresa. Embora eles fossem capazes de demonstrar que a linhagem celular 2BN era deficiente para reparo de quebras bifilamentares, eles não foram capazes de restaurar o sistema de reparo e produzir o fenótipo selvagem por complementação genética com nenhum dos genes que codificam as proteínas NHEJ. Isto é, quando eles introduziram genes tipo selvagem codificando proteínas NHEJ (por exemplo, KU70, KU80, ligase IV) na linhagem 2BN, a divisão celular ainda era defeituosa no reparo de quebras bifilamentares. Esse resultado negativo indicou que a linhagem celular 2BN tinha uma mutação em uma proteína NHEJ desconhecida.

Na era da genômica, a identificação de proteínas ligadas a doenças está se tornando mais comum devido à grande disponibilidade de linhagens celulares de pessoas com fenótipos de doenças. Quando nós, humanos, temos um problema de saúde, vamos a um médico e lhe contamos nossos sintomas, incluindo informações sobre parentes com problemas similares. Tal informação é de crescente impor-

tância na era da genômica, com suas sempre crescentes ferramentas genéticas que, geralmente, podem ser usadas para identificar genes mutantes associados a distúrbios herdados (veja os Capítulos 13 e 20).

O que os geneticistas podem fazer no laboratório para encontrar uma proteína, tal como a desconhecida proteína NHEJ, que ainda não foi identificada? Dois laboratórios usando enfoques muito diferentes foram bem-sucedidos em identificar o novo componente NHEJ. Um enfoque será descrito aqui porque foi empregado com sucesso para descobrir várias outras proteínas. Como notado nos capítulos anteriores, muitas proteínas celulares fazem seu trabalho interagindo com outras proteínas. O Capítulo 13, por exemplo, descreveu o teste diíbrido de leveduras usado para identificar proteínas que interagem com a proteína de interesse. No caso agora em consideração, a proteína de interesse era o componente XRCC4 de NHEJ (veja Figura 15.31), e o teste diíbrido identificou uma proteína interagindo, de 33 kD, que era codificada por uma matriz de leitura aberta humana não-caracterizada. Se duas proteínas interagem no teste de levedura diíbrida, isso não significa, necessariamente, que essas proteínas interagem nas células humanas. Para estabelecer uma ligação entre a proteína de 33 kD e a via NHEJ, os geneticistas usaram outra técnica valiosa de suas ferramentas, RNAi (veja os Capítulos 8 e 13). Nesse caso, eles demonstraram que as células normais expressando RNA anti-sentido da ORF que codifica a proteína de 33 kD, que evitaria a tradução desse gene em proteína, seriam hoje defectivas na execução da via NHEJ.

Essa história se completou ao se mostrar que as células 2BN defectivas no reparo bifilamentar não têm a proteína de 33 kD. A expressão dessa proteína corrige os defeitos celulares.

NHEJ é uma via propensa a erro que repara quebras bifilamentares em eucariontes superiores ligando as pontas livres de volta. A identificação dos genes responsáveis pelos distúrbios herdados é um modo importante usado pelos geneticistas para isolar componentes antes desconhecidos de reparo e outras vias biológicas.

Recombinação homóloga Se uma quebra bifilamentar ocorre após a replicação de uma região cromossômica em uma célula em divisão, o dano pode ser corrigido por um mecanismo livre de erro chamado de **helicoidização de filamento dependente de síntese (SDSA)**. Esse mecanismo é mostrado na Figura 15.30. Ele usa cromátides-irmãs disponíveis na mitose como moldes para garantir o reparo correto.

As primeiras etapas na SDSA são a ligação das pontas quebradas por proteínas especializadas e enzimas, o aparo das pontas 5' por uma endonuclease para expor regiões unifilamentares e o revestimento dessas regiões com proteínas que incluem o homólogo de RecA, Rad51. Lembre que, na resposta SOS, os monômeros RecA associam-se com regiões de DNA unifilamentar para formar filamentos de nucleoproteína. Similarmente, Rad51 forma longos filamentos à medida que se associa com a região unifilamentar exposta. O filamento de DNA-Rad51 então toma parte de pesquisa marcante da cromátide-irmã não-danificada quanto à sequência complementar que será usada como molde

Reparo livre de erro de quebras bifilamentares por SDSA

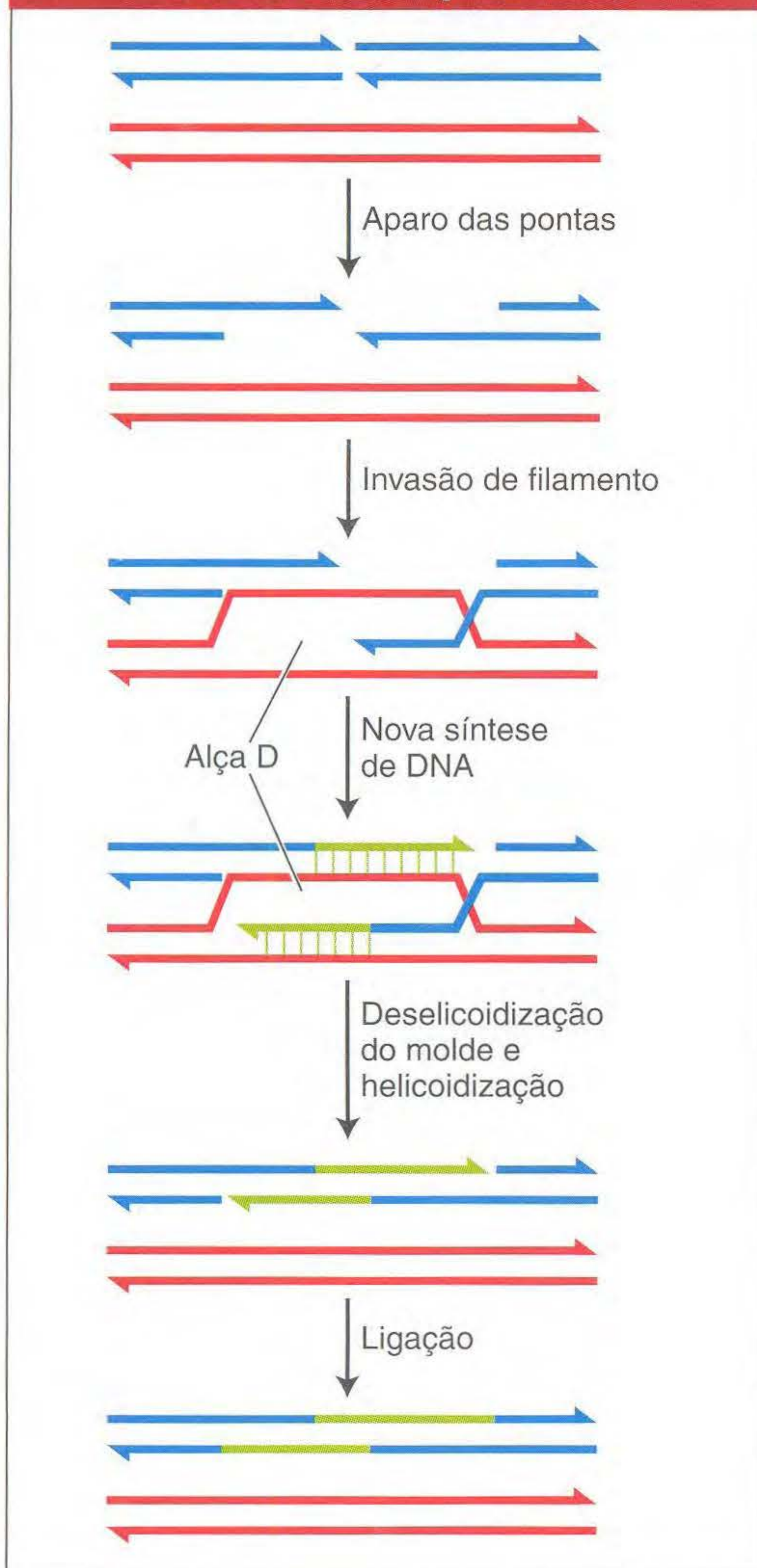


FIG. 15.30 O mecanismo livre de erro de helicoidização de filamento dependente de síntese (SDSA) repara quebras bifilamentares nas células em divisão.

para a síntese de DNA. Esse processo é chamado de invasão de filamento. A ponta 3' do filamento invasor desloca uma das cromátides-irmãs não-danificadas, o que forma uma alça D (de deslocamento), e inicia a síntese de DNA a partir de sua ponta 3' livre. A nova síntese de DNA continua a partir de ambas as pontas 3' até que ambos os filamentos se desenrolem de seus moldes e se helicoidizem. A ligação fecha os cortes, deixando um trecho reparado de DNA que tem uma característica muito distinta: ele foi replicado por um processo conservativo. Isto é, ambos os filamentos são recém-sintetizados, em acentuado contraste à replicação semiconservativa da maioria do DNA (veja o Capítulo 7).

A helicoidização de filamento dependente de síntese é um mecanismo livre de erro que repara quebras bifilamentares nas células em divisão nas quais uma cromátide-irmã está disponível para servir como molde para a síntese de reparo.

► 15.5 O Mecanismo de Crossing Over Meiótico

Nossa consideração sobre o reparo de quebras bifilamentares nas células em divisão leva naturalmente ao tópico de crossing over na meiose porque uma quebra bifilamentar inicia o evento do crossing. Embora as quebras sejam uma parte normal e essencial da meiose, elas são, se não processadas correta e eficientemente, tão perigosas quanto as quebras acidentais já discutidas.

Embora o processo da meiose tenha sido primeiro descrito há mais de um século, o mecanismo subjacente à recombinação meiótica ainda é uma área ativa de pesquisa que produziu importantes descobertas dentro dos últimos 10 anos.

Quebras bifilamentares programadas iniciam a recombinação meiótica

O crossing over é um processo marcadamente preciso que ocorre entre dois cromossomos homólogos (Figura 15.31). Os primeiros modelos propuseram que quebras unifilamentares em cromátides não-irmãs iniciavam a recombinação. Entretanto, a recombinação hoje é conhecida como sendo iniciada quando uma enzima chamada SpoII faz cortes bifilamentares no DNA em um dos cromossomos homólogos após a forquilha de replicação ter passado por uma região cromossômica (Figura 15.32). Embora primeiro descoberta em leveduras, a proteína SpoII é amplamente conservada em eucariontes, indicando que esse mecanismo para iniciar a recombinação é amplamente empregado.

Após fazer seus cortes, a enzima SpoII permanece ligada à agora ponta 5' livre, onde parece atender a duas finalidades: primeira, ela protege as pontas de outros danos, incluindo a recombinação espúria com outras pontas livres; segunda, ela pode atrair outras proteínas que são necessárias para a etapa seguinte na recombinação. Essa etapa é realmente muito similar à que acontece no sistema de reparo de quebras bifilamentares nas células em divisão. As pontas 5' são aparadas e um complexo proteico liga-se às pontas 3' unifilamentares (veja Figura 15.30).

Crossing durante a meiose

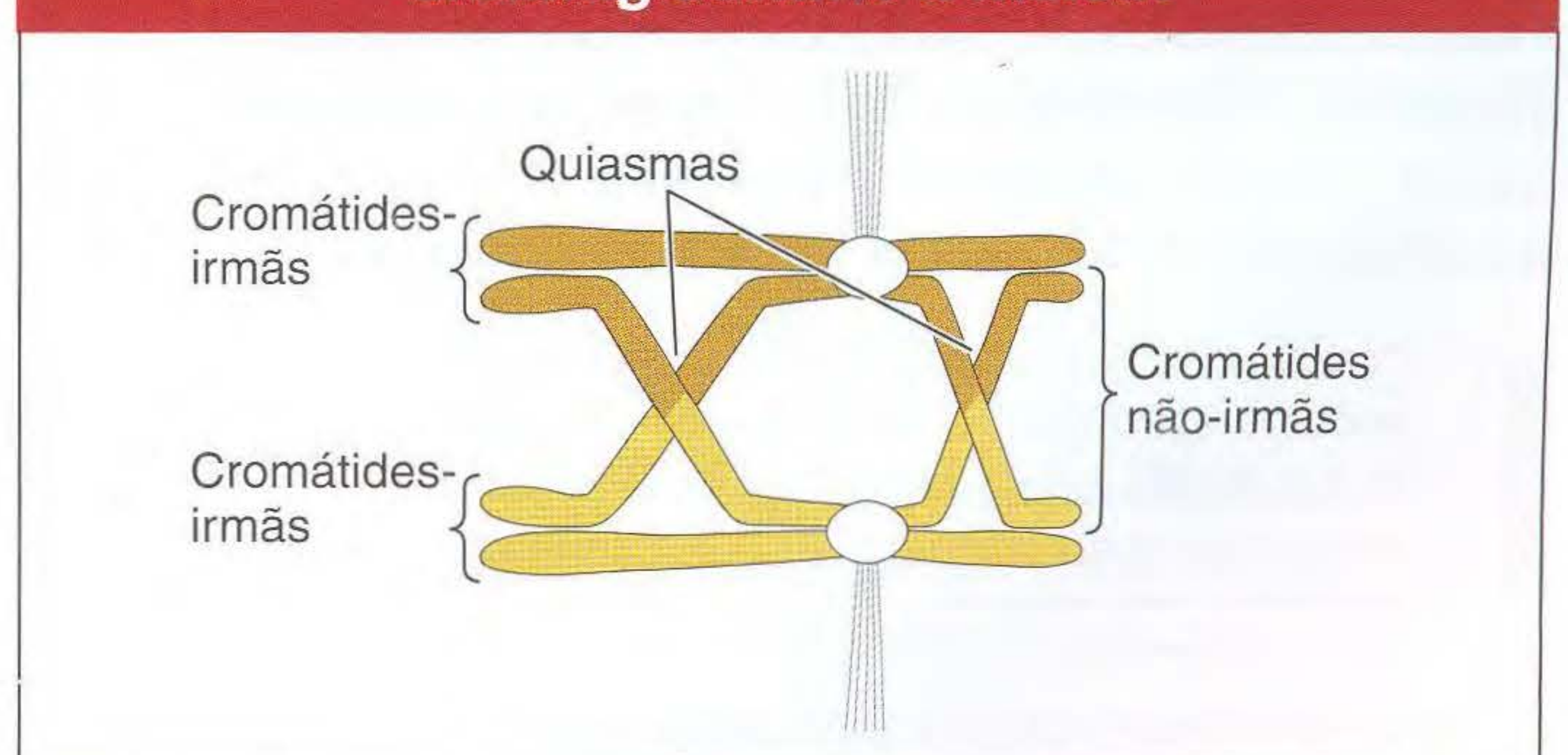


FIG. 15.31 Troca de braços cromossômicos entre cromátides não-irmãs durante a meiose produz um quiasma, a localização dos crossings. Os círculos representam os centrômeros que estão ligados às fibras do fuso. [Adaptado de Mathew J. Neale and Scott Keeney. *Nature* 442, 2006, 153–158.]

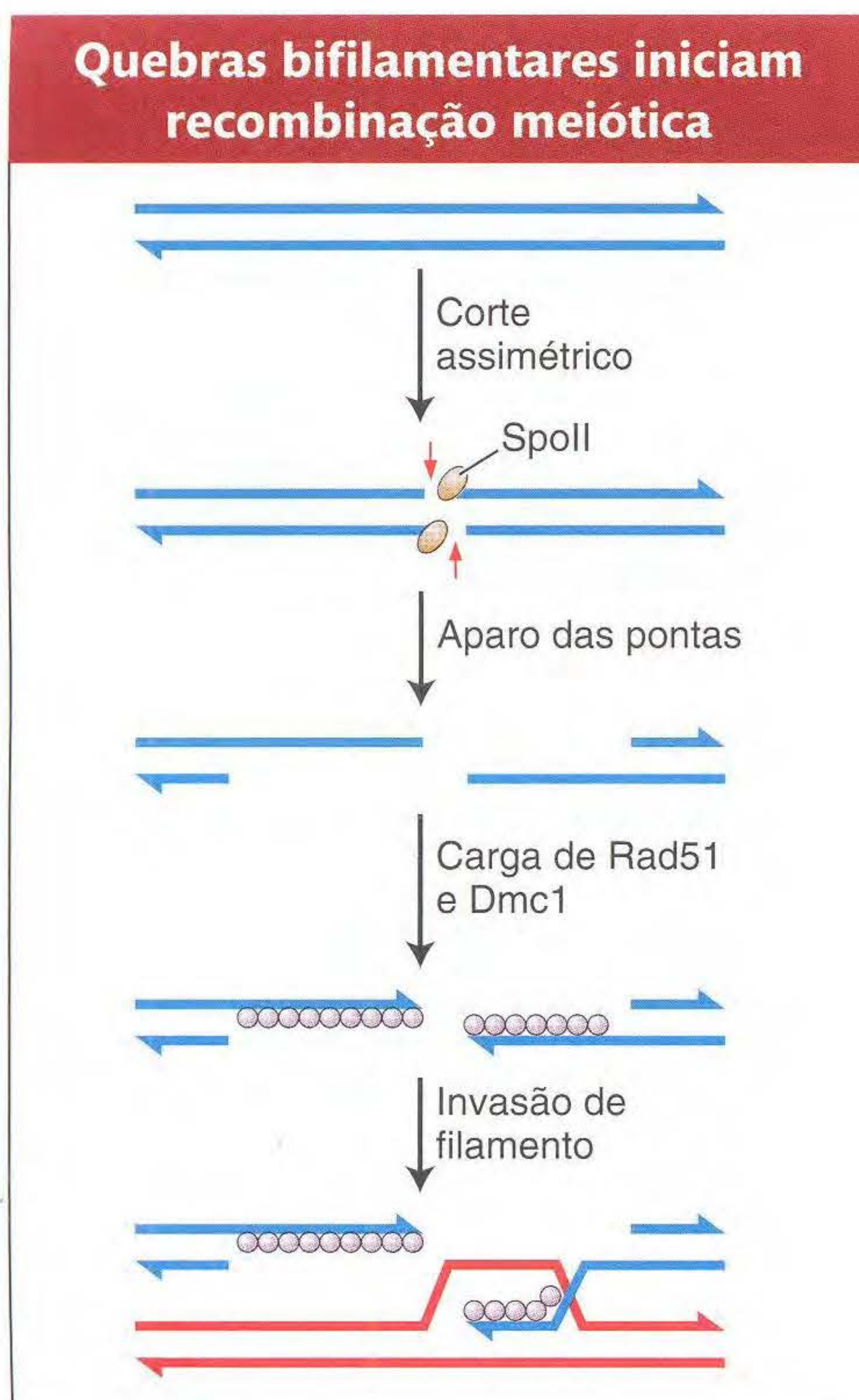


FIG. 15.32 A recombinação meiótica é iniciada quando a enzima SpoII faz cortes desencontrados em um par de filamentos de DNA em uma cromátide.

Esse complexo inclui a proteína Rad51, que, como, já mencionado, é homóloga à proteína RecA que toma parte nessa marcante procura de complementariedade na cromátide-irmã.

Nessa época, a recombinação meiótica toma um caminho marcadamente diferente do reparo de quebra bifilamentar. Na meiose, Rad51 associa-se a outra proteína, Dmc1, que está presente apenas durante a meiose (veja Figura 15.32). (Deve ser notado que os organismos-modelo *Drosophila* e *C. elegans* não têm homólogos Dmc1.) De algum modo, por um mecanismo não de todo compreendido, o filamento contendo Rad51-Dmc1 conduz a uma procura de uma sequência complementar. Entretanto, em contraste com o reparo de quebras bifilamentares, o filamento procura o cromossomo homólogo, não a cromátide-irmã. A pesquisa culmina na invasão filamental e formação de alça D, como no reparo da quebra bifilamentar. Esses eventos são necessários para a formação de quiasmas na meiose I. Isto é, os homólogos tornam-se conectados como resultado da recombinação.

A recombinação meiótica é iniciada pela enzima SpoII, que introduz cortes bifilamentares nos cromossomos após eles terem se replicado, mas antes da separação dos homólogos.

Esse processo é direto até atingir esse ponto. Entretanto, deixamos convenientemente de fora dois fatores complicantes. O primeiro fator é que os quiasmas conectando cromossomos homólogos são bem menos numerosos que o número de cortes feitos por SpoII. Existe pelo menos um quiasma

por cromossomo (alguns organismos, tais como *Arabidopsis* e leveduras, têm mais). Em contraste, mais de 200 quebras bifilamentares são introduzidas nos 16 cromossomos de leveduras, e provavelmente muitas mais em eucariontes superiores. A totalidade ou a maioria dessas quebras é aparada, de modo que, em média, cerca de 1 kb de DNA unifilamentar é exposto em cada quebra. A totalidade ou a maior parte delas toma parte nas pesquisas Rad51-Dmc1, com a subsequente formação de alças D mostradas na Figura 15.32. A primeira complicação é que a grande maioria dessas quebras não leva a quiasmas. O que ocorre então?

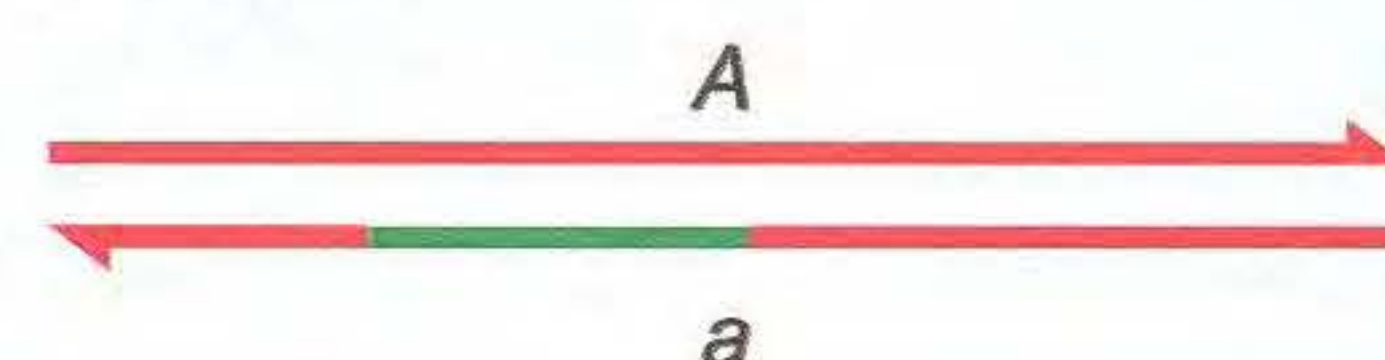
O segundo fator complicador é que o crossing over meiótico na verdade tem resultados genéticos possivelmente diferentes. No restante deste capítulo, veremos quais são os resultados e como o atual modelo explica tanto os resultados genéticos quanto o destino das quebras bifilamentares.

A análise genética das tétrades fornece indícios do mecanismo de recombinação

A análise de tétrades em fungo filamentos, tal como *Neurospora crassa*, deu o impulso para os primeiros modelos de recombinação meiótica. Como dito no Capítulo 3, a vantagem de estudar fungos é que todos os quatro produtos de uma única meiose podem ser recuperados e examinados. Esperamos que um cruzamento $A \times a$ em fungos crie um meiócito híbrido A/a que segrega em uma proporção 1:1 para formar os produtos meióticos de acordo com a primeira lei de Mendel de segregação igual. De fato, a proporção 1:1 é encontrada na maioria dos meiócitos de fungos: vemos 4 A e 4 a . Entretanto, em raros meiócitos, qualquer um dos quatro tipos de proporções aberrantes pode ser encontrado, e esses resultados genéticos diferentes dão os indícios necessários para construir um modelo de crossing.

A Figura 15.33 mostra as proporções aberrantes mais comumente obtidas. Parece que alguns alelos no cruzamento foram “convertidos” nos alelos opostos. Esse processo ficou conhecido como **conversão gênica**. Ele só pode ser detectado onde há heterozigose para dois alelos diferentes de um gene. Nos ascas com uma proporção 6:2 ou 2:6, ambos os filamentos de uma cromátide parecem ter sido convertidos. Nos ascos com uma proporção 5:3 ou 3:5, apenas um filamento de uma cromátide de dois filamentos parece ter sido convertido. Para ver por que, observe na Figura 15.33 que, nos ascos com proporção 5:3 ou 3:5, membros diferentes de um par de esporos têm genótipos diferentes. Lembre que cada par de esporos é produzido por mitose de um único produto de meiose. As proporções 5:3 ou 3:5 podem ser explicadas apenas se um filamento da dupla hélice levar informações de um alelo diferente de um gene A na conclusão da meiose.

Assim, a conversão gênica pode ser explicada se o processo da meiose formar um DNA heterodúplex — isto é, DNA com um segmento no qual um filamento é a sequência de nucleotídeos do alelo A e um filamento é a sequência de nucleotídeos do alelo a .



A conversão gênica resulta em proporções aberrantes de segregação

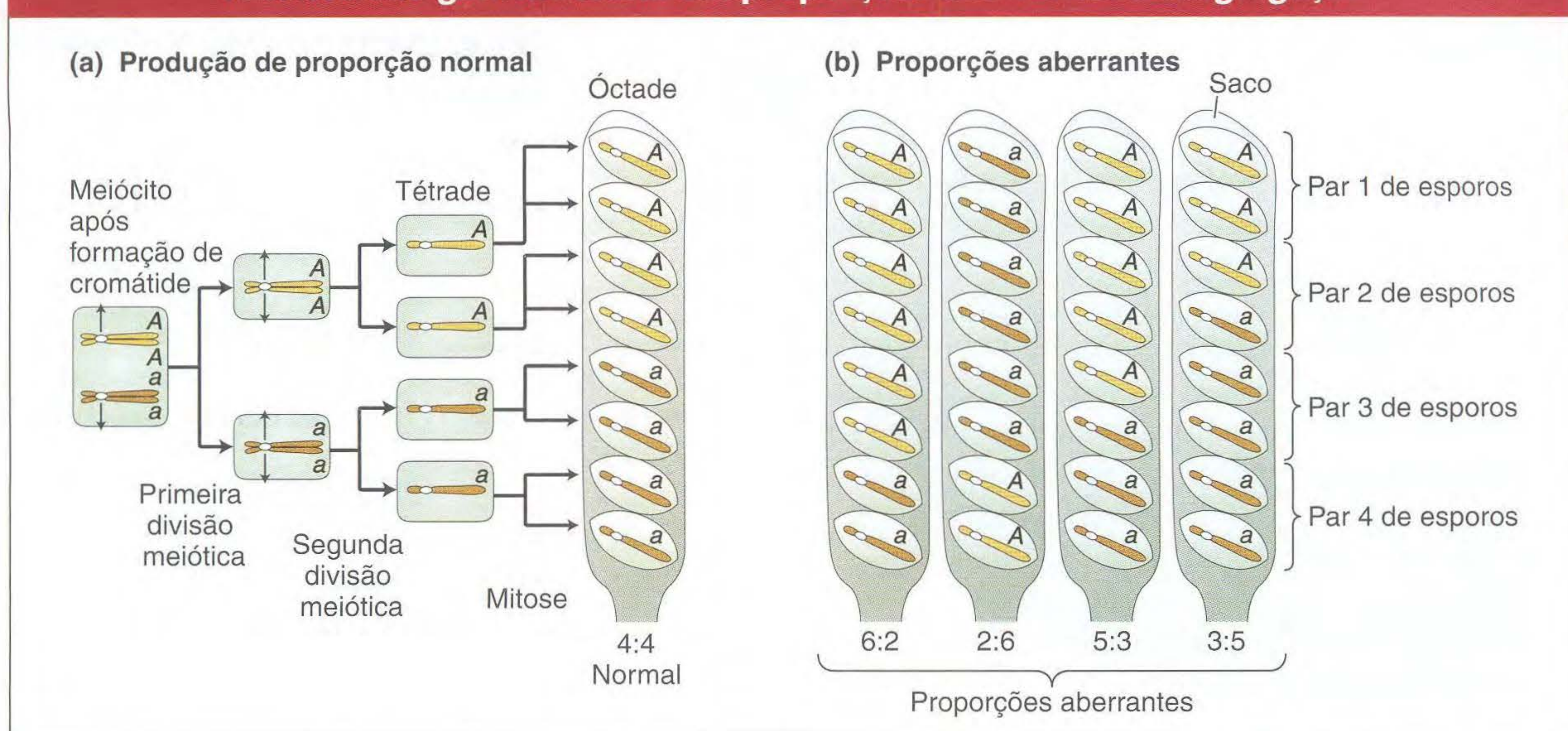
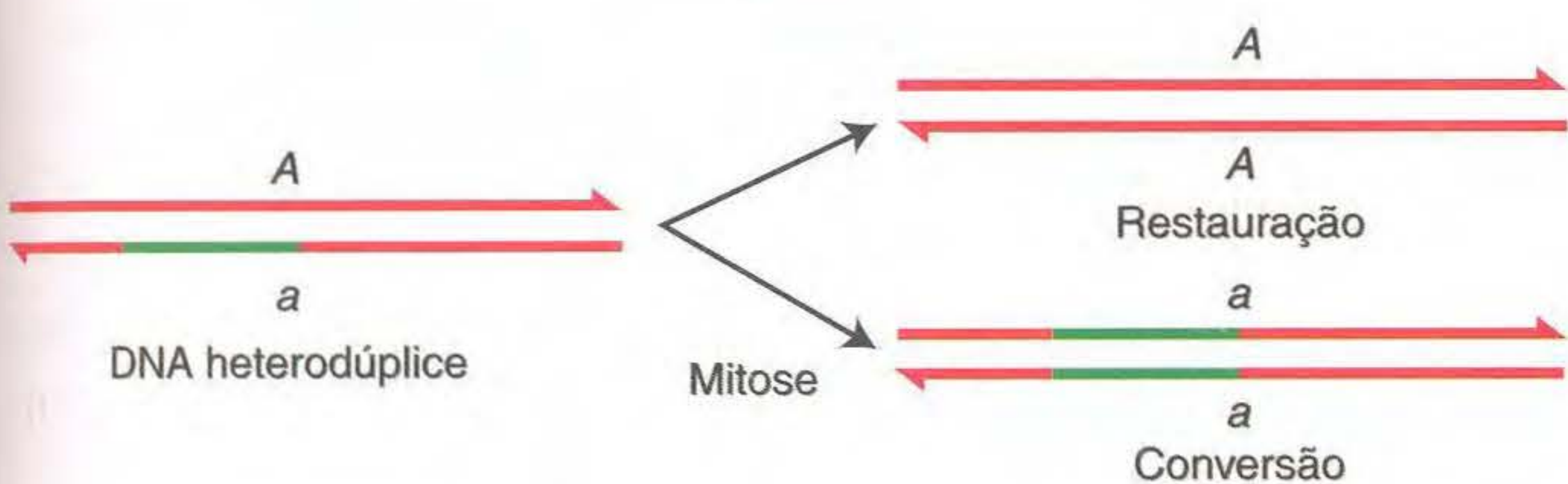


FIG. 15.33 Segregação aberrante em *Neurospora*. (a) Normalmente, um meiócito *A/a* sofre meiose e, então, mitose, resultando em uma proporção de 4:4 de produtos *A* e *a*. (b) Raramente, a conversão gênica resulta em proporções aberrantes.

Após a mitose, as duas cromátides-irmãs resultantes da divisão de tal heterodúplce contendo núcleo serão diferentes, uma com *A* e outra com *a*.



Suponhamos que *A* e *a* difiram por um único par de bases; que o par de bases seja G-C em *A* e A-T em *a*. Vamos supor ainda que, para qualquer determinado gene, um raro heterodúplce único seja formado na meiose de modo que A-T em um dos produtos se torne G-T (logo veremos o mecanismo). Então, podemos representar os produtos da meiose como

1. G-C 2. G-C 3. **G-T** 4. A-T

Após a mitose pós-meiótica, a óctade resultante será

1. G-C 2. G-C 3. G-C 4. G-C 5. **G-C** 6. **A-T**
7. A-T 8. A-T

que é a proporção 5:3 observada.

Entretanto, pelo que aprendemos neste capítulo, vemos que G-T do heterodúplce é um candidato a pareamento errado. Suponhamos que tal sistema de reparo remova o T do heterodúplce G-T e insira C, e assim tenhamos G-C. Nes-

se caso, haverá uma proporção de 3:1 de GC:AT e a óctade apresentará uma proporção de 6:2.

O modelo de quebra bifilamentar para recombinação meiótica

Um modelo para explicar a origem do DNA heterodúplce em uma recombinação meiótica é mostrado na Figura 15.34, à esquerda. Começaremos no ponto da invasão do filamento (onde paramos na Figura 15.32). O invasor inicia a síntese usando o filamento complementar da cromátide *A* como molde. Essa nova síntese aumenta o filamento único alça D, que serve como molde para a síntese de DNA iniciada pelo outro (*a*) filamento único. O preenchimento dos espaços pela atividade da polimerase e a junção das pontas de DNA pela ligase resultam em uma estrutura peculiar que parece dois crossings unifilamentares. Note que essa estrutura também contém o heterodúplce de que precisamos.

Os “crossings” unifilamentares são chamados de **junções de Holliday**, em homenagem a Robin Holliday, que primeiro as propôs nos anos de 1960. As duas junções são chamadas de **junções duplas de Holliday**. O modelo propõe que essas estruturas instáveis são resolvidas para produzir um crossing recíproco bifilamentar. Além disso, o modelo explica as ligações físicas importantes entre cromossomos homólogos na meiose. As duplas junções de Holliday são instáveis e devem ser resolvidas de um dentre dois modos. De modo simples, elas podem ser resolvidas por quebras “verticalmente” ou “horizontalmente” e reunião de filamentos únicos (veja setas na Figura 15.36). Uma resolução resulta em um crossing recíproco bifilamentar (mostrado à esquerda) e a outra um não-crossing (direita).

O modelo de decisão precoce de crossing

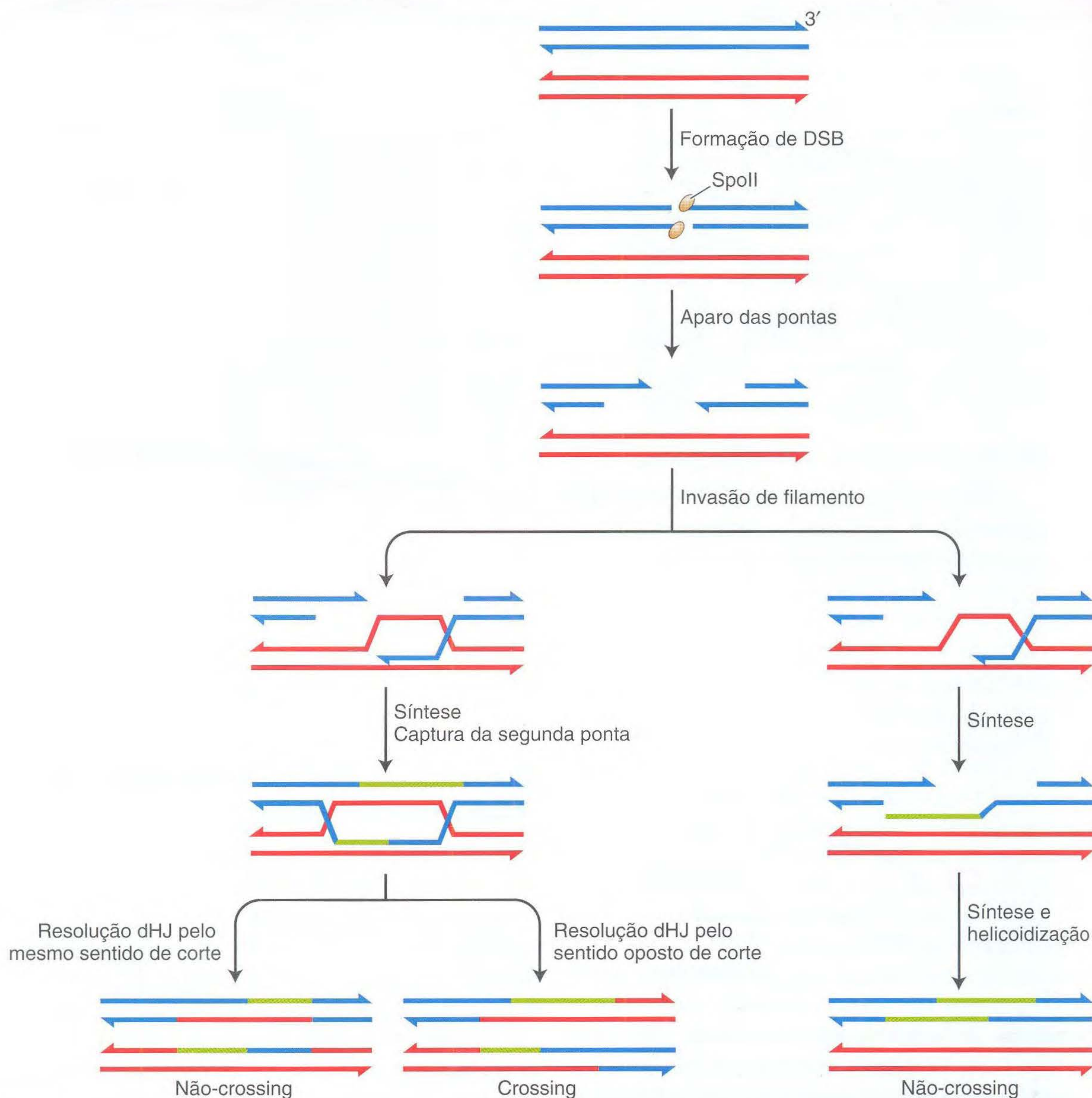


FIG. 15.34 De acordo com o modelo de decisão precoce de crossing, surgem não-crossings e crossings por duas vias distintas. [Adaptado da Figura 2 de Mathew J. Neale and Scott Keeney, *Nature* 442, 2006, 153–158.]

Esse modelo resolve um fator complicante, os diferentes resultados genéticos. Mas o modelo também deve explicar o fato de que, em leveduras, parece haver muito mais cortes do que nas junções duplas de Holliday e quiasmas. Assim, a maioria dos cortes resulta em não-crossings. A que ponto no processo são tomadas as decisões de seguir um caminho ou outro? Os resultados de novos experimentos sugerem que, em leveduras, a decisão é tomada após SpoII cortar, mas antes da invasão e formação da alça D.

Esse achado e outros novos resultados levaram ao modelo mostrado com a terceira via na Figura 15.34. O **modelo de decisão precoce de crossing** diz que duas vias distintas levam a não-crossings ou crossings. Assim, a maioria dos cortes duplos são resolvidos como não-crossings por um mecanismo que é similar

à helicoidização de filamento dependente de síntese (veja Figura 15.34), à exceção de que a reação é entre cromátides não-irmãs. A via do crossing supostamente inclui um conjunto de proteínas que estabilizam a dupla junção de Holliday, que é normalmente uma estrutura inerentemente instável. Esse modelo prevê que a mutação dessas proteínas impedirá crossings, mas não terá efeito nos não-crossings. Tais mutações de fato foram isoladas.

Os produtos de crossing e não-crossing da recombinação meiótica surgem de cortes bifilamentares programados. Os crossings são formados na resolução de junções duplas de Holliday, enquanto a via que leva a não-crossings pode envolver uma resolução alternativa das junções duplas de Holliday ou pode ser distinta e assemelhar-se à via de reparo SDSA.

15.6 Câncer: Uma Consequência Fenotípica Importante da Mutação



Por que tantos agentes mutagênicos causam câncer? Qual a ligação entre câncer e mutação? Nesta seção, exploramos a ligação câncer-mutação.

Ficou claro que praticamente todos os cânceres de células somáticas surgem devido a uma série de mutações específicas que se acumulam em uma célula. Algumas dessas mutações alteram a atividade de um gene; outras simplesmente eliminam a atividade gênica. As mutações promotoras de câncer são enquadradas em uma dentre algumas categorias importantes: as que aumentam a habilidade da célula em proliferar; as que diminuem a suscetibilidade de uma célula à via de suicídio, chamada de **apoptose**; ou as que aumentam a taxa geral de mutações da célula ou sua longevidade, de modo que todas as mutações, inclusive as que estimulam a proliferação ou apoptose, são mais prováveis de ocorrer.

Como as células cancerosas diferem das células normais

Um tumor maligno, ou **câncer**, é um agregado de células, todas descendentes de uma célula inicial fundadora aberrante. Em outras palavras, as células malignas são todas membros de um único clone, mesmo nos cânceres avançados que têm muitos tumores em muitos locais no corpo. As células cancerosas tipicamente diferem de suas vizinhas normais por várias características fenotípicas, tais como rápida taxa de divisão, habilidade de invadir novos territórios celulares, alta taxa metabólica e forma anormal. Por exemplo, quando as células de camadas celulares epiteliais normais são colocadas em cultura, elas só podem crescer quando ancoradas a um

frasco de cultura. Além disso, as células epiteliais normais em cultura só se dividem até formarem uma única camada contínua (Figura 15.35a). Nesse ponto, elas de algum modo reconhecem que formaram uma única camada epitelial e param de se dividir. Em contraste, as células malignas derivadas de tecido epitelial continuam a proliferar, empilhando-se umas sobre as outras (Figura 15.35b).

Claramente, os fatores que regulam a fisiologia celular normal foram alterados. Qual então é a causa subjacente do câncer? Muitos tipos celulares diferentes causam câncer? Muitos tipos celulares diferentes podem ser convertidos em um estado maligno? Há um tema comum? Ou cada um surge de um modo diferente? Podemos pensar sobre o câncer, de um modo geral, como sendo devido ao acúmulo de múltiplas mutações em uma única célula que faz com que ela prolifere fora de controle. Algumas dessas mutações podem ser transmitidas pelos genitores através da linhagem germinativa. A maior parte surge *de novo* na linhagem somática de determinada célula.

Mutações em células cancerosas

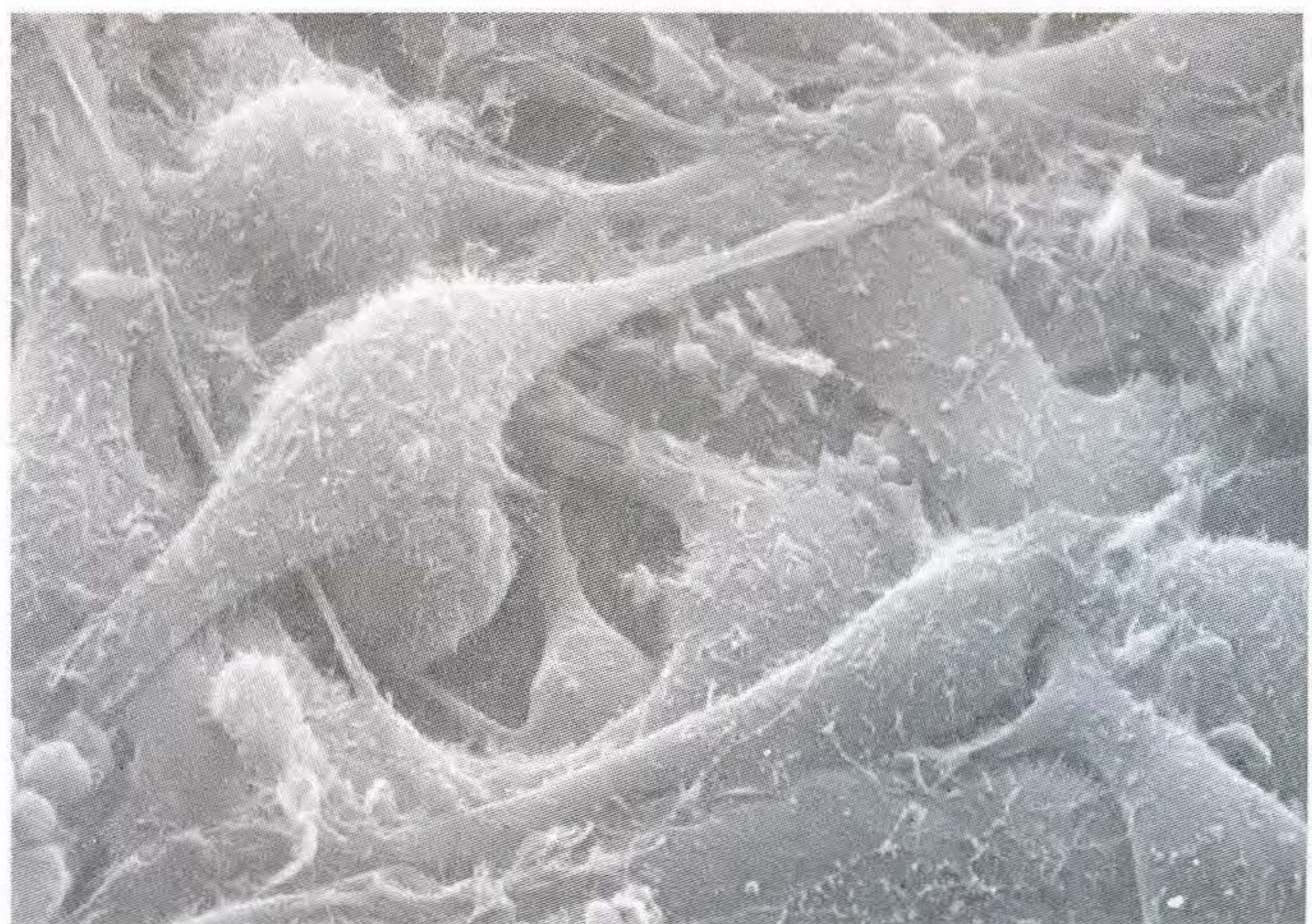
Várias linhas de evidência indicam uma origem genética para a transformação de células do estado benigno para o canceroso. Primeiro, como já discutido neste capítulo, muitos agentes mutagênicos, tais como substâncias químicas e radiação, causam câncer, sugerindo que produzem câncer introduzindo mutações nos genes. Segundo, e mais importante, as mutações que estão freqüentemente associadas a tipos particulares de cânceres já foram identificadas.

Dois tipos gerais estão associados a tumores: mutações oncogênicas e mutações em genes supressores tumorais. As mutações **oncogênicas** atuam em uma célula cancerosa como mutações dominantes de ganho de função (veja o Capítulo 6 para uma discussão de mutações dominantes de

Células normais e células transformadas por um oncogene



(a)



(b)

FIG. 15.35 Micrografias eletrônicas de varredura de (a) células normais e (b) células transformadas pelo vírus do sarcoma de Rous, que infecta células com o oncogene *src*. (a) Uma linhagem celular normal chamada de 3T3. Note a estrutura organizada de monocamada das células. (b) Uma derivativa transformada de 3T3. Note como as células são mais arredondadas e empilhadas umas sobre as outras. [Cortesia de L.-B. Chen.]

ganho de função). Essa afirmativa sugere duas características principais de mutações oncogênicas: primeira, as proteínas codificadas por oncogenes são geralmente *ativadas* em células tumorais; e, segunda, a mutação precisa estar presente apenas em um alelo para contribuir para a formação do tumor. O gene em sua forma normal, não-mutada, é chamado de **proto-oncogene**.

As mutações nos **genes supressores tumorais** que promovem a formação de tumor são mutações recessivas de perda de função. Isto é, esse tipo de mutação faz com que os produtos dos genes codificados percam muito da sua atividade ou toda ela (isto é, a mutação é uma mutação nula). Além disso, para que o câncer se desenvolva, a mutação deve estar presente em ambos os alelos do gene.

Os oncogenes codificam formas mutadas de proteínas celulares normais que resultam em mutações dominantes, geralmente devido à sua inativação imprópria. Em contraste, os genes supressores tumorais codificam proteínas cuja perda de atividade pode contribuir para um estado canceroso. Assim sendo, são mutações recessivas.

Classes de oncogenes Já foram identificados cerca de 100 oncogenes diferentes. Como suas contrapartes normais — os proto-oncogenes — funcionam? Os proto-oncogenes geralmente codificam uma classe de proteínas que são ativas apenas quando os sinais de regulação apropriados permitem que eles sejam ativados. Muitos produtos proto-oncogênicos são elementos em vias que induzem (controle positivo) o ciclo celular. Esses produtos incluem receptores de fatores de crescimento, proteínas de transdução de sinal e reguladores transcricionais. Outros produtos de proto-oncogenes atuam inibindo (controle negativo) a via apoptótica que destrói células danificadas. Em ambos os tipos de mutação de oncogene, a

atividade da proteína mutante foi desacoplada de sua via reguladora normal, levando à sua expressão desregulada contínua. O produto proteico expresso continuamente por um oncogene é chamado de **oncoproteína**. Várias categorias de oncogenes foram identificadas de acordo com os modos diferentes nos quais as funções regulatórias foram desacopladas.

O oncogene *ras* pode ser usado para ilustrar o que ocorre quando um gene normal mantém uma mutação promotora de tumor. Como geralmente é o caso, a mutação de uma proteína normal para uma oncoproteína envolve modificações estruturais na própria proteína, nesse caso causada por uma simples mutação de ponto. Uma única substituição de par de bases que converte glicina em valina no aminoácido 12 da proteína Ras, por exemplo, cria a oncoproteína encontrada no câncer de bexiga humana (Figura 15.36a). A proteína Ras normal é uma subunidade de proteína G que toma parte na transdução de sinal. Ela normalmente funciona oscilando entre o estado ativo ligado a GTP e o inativo ligado a GDP. A mutação de sentido trocado no oncogene *ras* produz uma oncoproteína que sempre liga GTP (Figura 15.36b), mesmo na ausência de sinais normais. Como consequência, a oncoproteína Ras propaga continuamente um sinal que promove a proliferação celular.

Genes supressores tumorais As funções normais dos genes supressores tumorais enquadram-se em categorias complementares às dos proto-oncogenes (veja Quadro 15.1). Alguns genes supressores tumorais codificam reguladores negativos cuja função normal é inibir o ciclo celular. Outros codificam reguladores positivos que normalmente ativam a apoptose, ou morte celular, de uma célula danificada. Outros ainda são participantes indiretos do câncer, com um papel normal no reparo de DNA danificado, ou no controle da longevidade celular. Consideraremos um exemplo aqui.

O oncogene *ras* é continuamente ativo

	Aminoácido 10			Aminoácido 15		
	Gli	Ala	Gli	Gli	Val	Gli
DNA tipo selvagem <i>ras</i>	GGC	GCC	GGC	GGT	GTG	GGC
DNA de oncogene <i>ras</i>	GGC	GCC	GTC	GGT	GTG	GGC
(a)			Val			

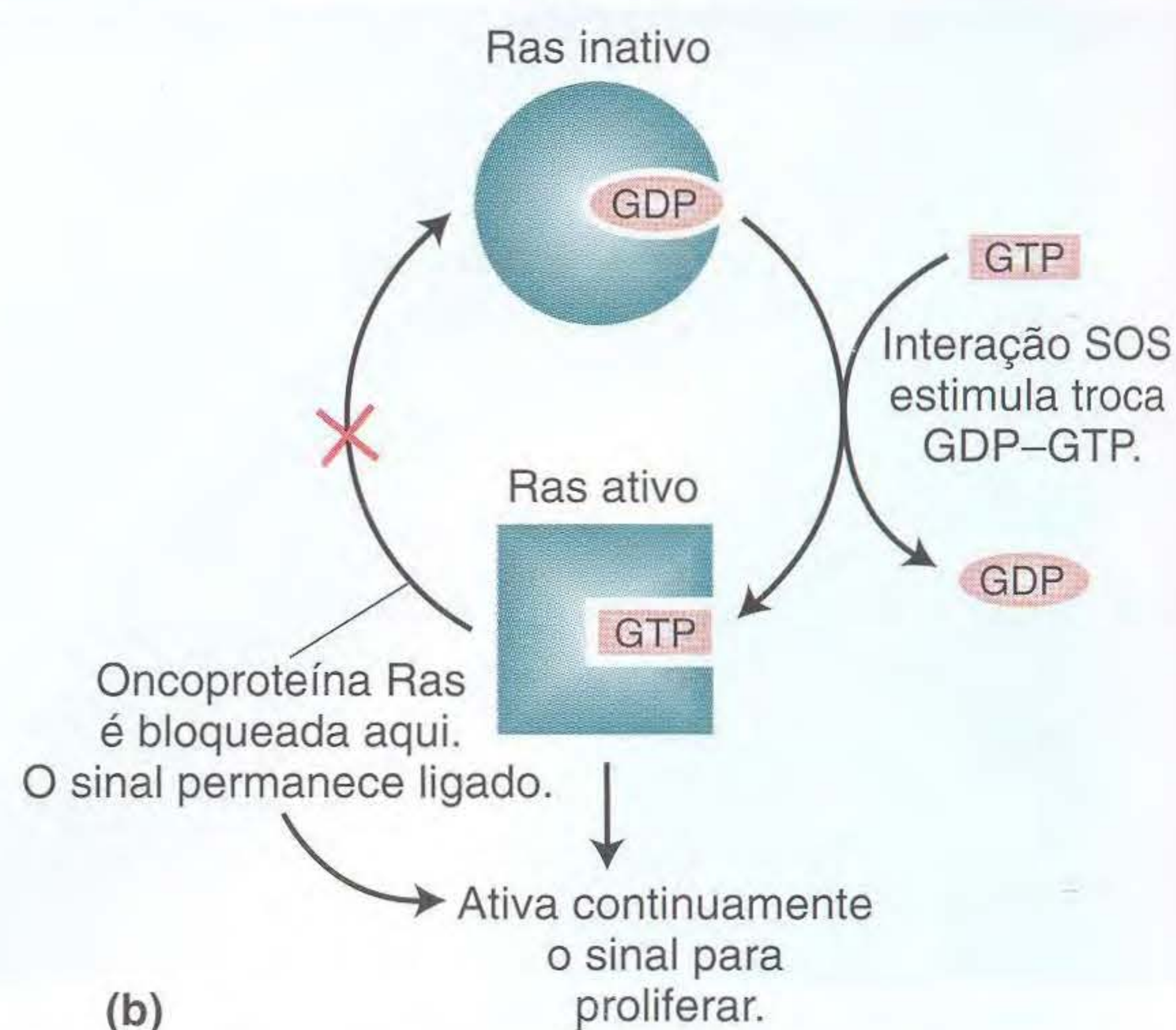


FIG. 15.36 Formação e efeito da oncoproteína Ras. (a) O oncogene *ras* difere do tipo selvagem por um único par de bases, produzindo uma oncoproteína Ras que difere do tipo selvagem em um aminoácido, na posição 12 na matriz de leitura aberta *ras*. (b) A oncoproteína Ras não pode hidrolisar GTP em GDP. Devido a esse defeito, a oncoproteína Ras permanece no complexo ativo Ras-GTP e ativa continuamente o sinal para proliferar.

Quadro 15.1 Funções de Proteínas Tipo Selvagem e Propriedades de Mutações Promotoras de Tumor nos Genes Correspondentes

Função da proteína tipo selvagem	Propriedades das mutações promotoras de tumor
Promove progressão do ciclo celular	Oncogene (ganho de função)
Inibe progressão do ciclo celular	Mutação de supressor tumoral (perda de função)
Promove apoptose	Mutação de supressor tumoral (perda de função)
Inibe apoptose	Oncogene (ganho de função)
Promove reparo de DNA	Mutação de supressor tumoral (perda de função)

As mutações no gene *p53* estão associadas a muitos tipos de tumores. De fato, estima-se que 50% dos tumores humanos não têm um gene *p53* funcional. A proteína *p53* ativa é um regulador transcricional que é ativado em resposta a dano ao DNA. A *p53* tipo selvagem ativada tem papel duplo: ela evita a progressão do ciclo celular até que o dano ao DNA seja reparado, e, sob algumas circunstâncias, induz apoptose. Se nenhum gene *p53* funcional está presente, o ciclo celular progride mesmo que o DNA danificado não tenha sido reparado. A progressão do ciclo celular para a mitose eleva a frequência geral de mutações, rearranjos cromossômicos e aneuploidia, e assim aumenta as chances de que surjam outras mutações que promovem a proliferação celular ou bloqueiam a apoptose.

Hoje está claro que as mutações que são capazes de elevar a taxa de mutação são importantes contribuintes para a progressão de tumores em humanos. Essas mutações são

recessivas nos genes supressores tumorais que normalmente funcionam nas vias de reparo do DNA. As mutações nesses genes interferem, assim, no reparo do DNA. Elas promovem indiretamente o crescimento tumoral elevando a taxa de mutação, o que torna muito mais provável que surja uma série de mutações de oncogenes e supressores tumorais, corrompendo a regulação normal do ciclo celular e a morte celular programada. Já foram identificadas muitas dessas mutações em genes supressores tumorais, incluindo algumas associadas a formas herdáveis de câncer em tecidos específicos. Os exemplos são as mutações *BRCA1* e *BRCA2* e o câncer de mama.

Os agentes mutagênicos podem causar alguns cânceres porque o câncer é, em parte, causado por versões mutantes de genes normais que levam a um crescimento descontrolado.

► **Resumo**

Uma mudança de DNA dentro de um gene (mutação de ponto) geralmente envolve uma ou algumas bases. As substituições de um só par de bases pode criar códons de sentido trocado ou sem sentido (término de transcrição). Uma purina substituída por outra purina (ou uma pirimidina substituída por outra pirimidina) é uma transição. Uma purina substituída por uma pirimidina (ou vice-versa) é uma transversão. A adição ou a deleção de um par de bases (indels) produz mudanças de matriz de leitura. Alguns genes humanos que contêm repetições de trinucleotídeos, especialmente as que são expressas no tecido neural, tornam-se mutadas pela expansão dessas repetições e podem, assim, causar doenças. A formação de repetições de monoaminoácidos dentro dos polipeptídeos codificados por esses genes é responsável pelos fenótipos mutantes.

As mutações podem ocorrer espontaneamente como subproduto de processos celulares normais, tais como replicação do DNA ou metabolismo, ou podem ser induzidas por radiação mutagênica ou substâncias químicas. Os mutágenos geralmente resultam em um tipo específico de mudança devido à sua especificidade química. Por exemplo, alguns produzem exclusivamente transições $G \cdot C \rightarrow A \cdot T$; outros, exclusivamente mudanças de matriz de leitura.

Embora as mutações sejam necessárias para gerar diversidade, muitas mutações estão associadas a doenças genéti-

cas herdadas, tais como xeroderma pigmentoso. Além disso, as mutações que ocorrem nas células somáticas são a fonte de muitos cânceres humanos. Muitas vias biológicas desenvolveram-se para corrigir o amplo espectro de mutações espontâneas e induzidas. Algumas vias, tais como o reparo por excisão de base e nucleotídeo e o reparo de mal pareamento, usam a informação inerente à complementariedade de bases para executar reparo livre de erro. Outras vias que usam polimerases de *bypass* para corrigir bases danificadas podem induzir erros na sequência de DNA.

A correção de quebras bifilamentares é particularmente importante porque essas lesões podem levar a rearranjos cromossômicos desestabilizantes. A junção de pontas não-homólogas é uma via que liga pontas quebradas, de modo que a forquilha de replicação parada não resulta na morte celular. Nas células em replicação, as quebras bifilamentares podem ser reparadas de um modo livre de erro pela via de helicoidização de filamento dependente de síntese, que usa a cromátide-irmã para reparar a quebra.

Centenas de quebras bifilamentares programadas iniciam o crossing meiótico entre cromátides não-irmãs. Como outras quebras bifilamentares, as quebras meióticas devem ser processadas rápida e eficientemente para evitar sérias consequências, tais como morte celular e câncer. Ainda está sendo estudado como esse reparo é feito. Um modelo atual,

informado pela segregação de características na formação de esporos de fungos, cita o uso de cromátides não-irmãs de cromossomos homólogos como molde para reparar as quebras. De acordo com esse modelo, a maioria das quebras leva a

produtos não-crossing e são provavelmente resolvidas por um mecanismo tipo SDSA. Entretanto, as quebras críticas para a formação de quiasmas levam a crossings seguindo a formação e resolução das junções duplas de Holliday.

► Termos-chave

agente intercalar	mudança tautomérica	reparo por excisão de base
análogo de base	mutação de ponto	reparo por excisão de nucleotídeo (NER)
apoptose	mutação de sentido trocado	reparo por excisão de nucleotídeo
câncer	mutação espontânea	acoplado à transcrição (TC-NER)
ceto	mutação indel	repetição de trinucleotídeo
conversão gênica	mutação induzida	síndrome de Cockayne
distúrbio genético heterogêneo	mutação sem sentido	síntese de DNA translesão
enol	mutação sinônima	sistema de reparo dependente de
gene supressor tumoral	mutagênese	homologia
helicoidização de filamento	mutágeno	sistema SOS
dependente de síntese (SDSA)	oncogene	sítio apurínico
imino	oncoproteína	substituição conservativa
junção de Holliday	plaqueamento réplica	substituição não-conservativa
junção de pontas não-homólogas	polimerase de <i>bypass</i> (translesão)	tautômero
(NHEJ)	polimerases translesão (<i>bypass</i>)	teste de Ames
junção dupla de Holliday	proto-oncogene	teste de flutuação
lesão espontânea	quebra bifilamentar	transição
modelo de decisão precoce de crossing	reparo de mal pareamento	transversão
mudança de matriz de leitura	reparo genômico global (GGR)	xeroderma pigmentoso

► Problemas Resolvidos

Problema resolvido 1. No Capítulo 9, aprendemos que os códons UAG e UAA são duas trincas sem sentido para o término de cadeia. Com base na especificidade da aflatoxina B₁ e etilmetanossulfonato (EMS), descreva se cada mutágeno seria capaz de reverter esses códons para o tipo selvagem.

SOLUÇÃO

EMS induz transições primariamente G · C → A · T. Os códons UAG não podem ser revertidos ao tipo selvagem, porque apenas a mudança UAG → UAA seria estimulada por EMS e geraria um códon sem sentido (ocre). Os códons UAA não seriam ativados por EMS. A aflatoxina B₁ induz primariamente transversões G · C → T · A. Apenas a terceira posição do códon UAG seria ativada, resultando em uma mudança UAG → UAU (ao nível de mRNA), que produz tirosina. Portanto, se tirosina fosse um aminoácido aceitável

no sítio correspondente na proteína, a aflatoxina B₁ poderia reverter os códons UAG. A aflatoxina B₁ não reverteria os códons UAA, pois nenhum par de bases G · C aparece na posição correspondente no DNA.

Problema resolvido 2. Explique por que as mutações induzidas por acridinas no fago T4 ou por ICR-191 em bactérias não podem ser revertidas por 5-bromouracil.

SOLUÇÃO

As acridinas e ICR-191 induzem mutações deletando ou adicionando um ou mais pares de bases, que resultam em uma mudança de matriz de leitura. Entretanto, 5-bromouracil induz mutações causando a substituição de uma base por outra. Essa substituição não pode compensar a mudança de matriz de leitura resultantes de ICR-191 e acridinas.

► Problemas

PROBLEMAS BÁSICOS

1. Considere as seguintes seqüências tipo selvagem e mutante:

Tipo selvagemCTTGCAAGCGAATC....

Mutante CTTGCTAGCGAATC....

- A substituição mostrada *parece* ter criado um códon de fim. Que outra informação você precisa para ter certeza que isso ocorreu?
2. Que tipo de mutação é representado pelas seguintes seqüências (mostradas como mRNA)?

Tipo selvagem5' AAUCCUUACGGA 3'....

Mutante5' AAUCCUACGGA 3'

3. Pode uma mutação de sentido trocado de prolina para histidina ser feita com mutágenos causadores de transição $G \cdot C \rightarrow A \cdot T$? E uma mutação de sentido trocado de prolina para serina?
4. Por substituição de par de bases, quais são todas as mudanças sinônimas que podem ser feitas começando com o códon CGG?
5. a. Quais são todas as transversões que podem ser feitas começando com o códon CGG?
b. Quais dessas transversões serão de sentido trocado? Você pode estar seguro?
6. Que tautômero de timina pode formar a maioria das pontes de hidrogênio no pareamento com outras bases do DNA?
7. a. Se a forma enol de timina é inserida em um molde unifilamentar no curso da replicação, que substituição de par de bases irá resultar?
b. Se a timina enoliza enquanto está agindo como um molde durante a replicação, que substituição de par de bases resultará?
8. a. A acridina laranja é um eficiente mutágeno para produzir alelos nulos por mutação. Por que produz alelos nulos?
b. Um certo composto similar à acridina gera apenas inserções únicas. Uma mutação induzida com esse composto é tratada com o mesmo composto, e são produzidos alguns revertentes. Como é possível esse resultado?
9. Defenda a frase: "Câncer é uma doença genética."
10. Cite um exemplo de um defeito de reparo de DNA que leve ao câncer.
11. No reparo de mal pareamento em *E. coli*, apenas um erro no filamento recém-sintetizado é corrigido. Como a *E. coli* é capaz de reconhecer o filamento recém-sintetizado? Por que essa habilidade faz sentido biológico?
12. Uma lesão mutacional resulta em uma sequência contendo um par de bases mal pareado:

5' AGCTGCCTT 3'
3' ACGATGGAA 5'
Códon

Se o reparo de mal pareamento ocorre em qualquer direção, que aminoácidos podem ser encontrados nesse sítio?

13. Sob que circunstâncias a junção de pontas não-homólogas pode ser considerada propensa a erro?
14. Por que muitas substâncias químicas que dão resultado positivo no teste de Ames também são classificadas como carcinogênicas?
15. A proteína SpoII é conservada em eucariontes. Você acha que ela também é conservada em espécies bacterianas? Justifique sua resposta.
16. Diferencie os elementos dos seguintes pares:
 - a. Transições e transversões
 - b. Mutações sinônimas e neutras
 - c. Mutações de sentido trocado e sem sentido
 - d. Mudança de matriz de leitura e mutações sem sentido

17. Quais as principais diferenças entre reparo de quebras bifilamentares quando elas são acidentais e seu reparo quando ocorrem na meiose?
18. O que é conversão gênica e o que isso tem a ver com reparo de mal pareamento?
19. Descreva duas lesões espontâneas que podem levar a mutações.
20. O que são polimerases de *bypass*? Como elas diferem das polimerases replicativas? Como suas características especiais facilitam seu papel no reparo do DNA?
21. O que é um heterodúplice de DNA? Qual a relação com a conversão gênica?
22. Nas células adultas que pararam de se dividir, que tipos de sistemas de reparo são possíveis?
23. Um certo composto que é um análogo da base citosina pode ser incorporado ao DNA. Ele normalmente faz pontes de hidrogênio como a citosina, mas, muito frequentemente, isomeriza-se em uma forma que faz pontes de hidrogênio como a timina. Você espera que esse composto seja mutagênico, e caso afirmativo, que tipos de mudança ele pode induzir no DNA?
24. Duas vias, recombinação homóloga e não-homóloga de junção de pontas (NHEJ), podem reparar quebras bifilamentares de DNA. Se a recombinação homóloga é uma via livre de erro enquanto NHEJ não é livre de erro, porque NHEJ é usada na maioria das vezes em eucariontes?
25. Que via de reparo reconhece danos ao DNA durante a transcrição? O que acontece se o dano não for reparado?

PROBLEMAS DESAFIADORES

26. a. Por que é impossível induzir mutações sem sentido (representadas a nível de mRNA pelas trincas UAG, UAA e UGA) tratando linhagens tipo selvagem com mutágenos que causam apenas transições $A \cdot T \rightarrow G \cdot C$ no DNA?
b. A hidroxilamina (HA) só causa transições $G \cdot C \rightarrow A \cdot T$ no DNA. A HA produz mutações em linhagens tipo selvagem?
c. O tratamento com HA reverte mutações sem sentido?
27. Várias mutações de ponto auxotróficas em *Neurospora* são tratadas com vários agentes para ver se ocorrerá reversão. Os seguintes resultados foram obtidos (um sinal mais indica reversão; HA só causa transições $G \cdot C \rightarrow A \cdot T$).

Mutante	5-BU	HA	Proflavina	Reversão espontânea
1	-	-	-	-
2	-	-	+	+
3	+	-	-	+
4	-	-	-	+
5	+	+	-	+

- a. Para cada um dos cinco mutantes, descreva a natureza da mutação original (não a reversão) a nível molecular. Seja o mais específico possível.



- b.** Para cada um dos cinco mutantes, cite um mutágeno possível que pode ter causado o evento mutacional original. (Mutação espontânea não é uma resposta aceitável.)
- c.** No experimento de reversão para o mutante 5, é obtido um derivativo prototrófico particularmente interessante. Quando esse tipo é cruzado com um tipo selvagem padrão, a prole consiste em 90% de prototróficos e 10% de auxotróficos. Explique esses resultados, incluindo o motivo exato para as frequências observadas.
- 28.** Você está usando nitrosoguanidina para “reverter” alelos *nic-2* mutantes (requer nicotinamida) em *Neurospora*. Você trata as células, plaqueia-as em um meio sem nicotinamida e procura colônias prototróficas. Você obtém os seguintes resultados para dois alelos mutantes. Explique estes resultados a nível molecular e indique como você testa suas hipóteses.
- a.** Com o alelo 1 *nic-2*, você não obtém nenhum prototrófico.
- b.** Com o alelo 2 *nic-2*, você obtém três colônias prototróficas A, B e C, e você cruza cada uma separadamente com a linhagem tipo selvagem. Do cruzamento prototrófico A × tipo selvagem, você obtém uma prole de 100, todos os quais são prototróficos. Do cruzamento prototrófico B × tipo selvagem, você obtém uma prole de 100, dos quais 78 são prototróficos e 22 requerem nicotinamida. Do cruzamento prototrófico C × tipo selvagem, você obtém uma prole de 1.000, dos quais 996 são prototróficos e 4 requerem nicotinamida.
- 29.** Você está trabalhando com um mutágeno recém-descoberto e quer determinar a mudança de bases que ele introduz no DNA. Você já determinou que o mutágeno altera quimicamente uma única base de tal modo que suas propriedades de pareamento de bases são alteradas permanentemente. Para determinar a especificidade da alteração, você examina as mudanças de aminoácidos que ocorrem após a mutagênese. Um exemplo do que você encontra é mostrado aqui:
Original: Gln-His-Ile-Glu-Lis
Mutante: Gln-His-Met-Glu-Lis
Original: Ala-Val-Asn-Arg
Mutante: Ala-Val-Ser-Arg
Original: Arg-Ser-Leu
Mutante: Arg-Ser-Leu-Trp-Lis-Tre-Fen
Qual a especificidade da mudança de base do mutágeno?
- 30.** Agora você encontra um mutante adicional do experimento no Problema 29:
Original: Ile-Leu-His-Gln
Mutante: Ile-Pro-His-Gln
Poderia a especificidade da mudança de base em sua resposta ao Problema 29 explicar essa mutação? Por que sim ou por que não?
- 31.** A linhagem A de *Neurospora* contém uma mutação *ad-3* que reverte espontaneamente a uma taxa de 10^{-26} . A linhagem A é cruzada com um isolado tipo selvagem recém-obtido, e as linhagens *ad-3* são recuperadas da prole. Quando a prole de 28 linhagens *ad-3* diferentes são

examinadas, são encontradas 13 linhagens que revertem a uma taxa de 1 em 10^{26} , mas as 15 linhagens restantes revertem à taxa de 1 em 10^{23} . Formule uma hipótese para explicar esses achados, e crie um programa experimental para testar sua hipótese.

- 32.** Você é um *expert* nos mecanismos de reparo de DNA. Você recebe uma amostra de uma linhagem celular humana derivada de uma mulher que tem sintomas de xeroderma pigmentoso. Você determina que ela tem uma mutação em um gene que não foi previamente associado a XP. Como isso é possível?
- 33.** Ozônio (O_3) é um componente importante de ocorrência natural em nossa atmosfera, onde forma uma camada que absorve radiação UV. Foi descoberto um buraco na camada de ozônio nos anos de 1970 sobre a Antártica e Austrália. O buraco aparece sazonalmente e foi considerado como devido à atividade humana. Especificamente, o ozônio é destruído por uma classe de substâncias (chamada de CFC para clorofluorocarbonos) que é encontrada em produtos de refrigeração, sistemas de ar condicionado e aerossol.
Como um cientista trabalhando em mecanismos de reparo de DNA, você descobre que há um aumento significativo de câncer de pele nas comunidades das praias da Austrália. Um repórter de jornal seu amigo lhe oferece publicar uma pequena nota (um parágrafo) no qual você descreve a possível ligação entre o buraco de ozônio e o aumento dos cânceres de pele. Com base no que você aprendeu sobre o reparo do DNA neste capítulo, escreva um parágrafo que explique a ligação entre o buraco e o câncer.
- 34.** Qual dos seguintes ascos lineares mostra conversão gênica no locus *arg-2*?

1	2	3	4	5	6
+	+	+	+	+	+
+	+	+	+	+	+
+	<i>arg</i>	+	+	<i>arg</i>	<i>arg</i>
+	<i>arg</i>	<i>arg</i>	<i>arg</i>	<i>arg</i>	<i>arg</i>
<i>arg</i>	<i>arg</i>	<i>arg</i>	+	+	<i>arg</i>
<i>arg</i>	<i>arg</i>	<i>arg</i>	<i>arg</i>	+	<i>arg</i>
<i>arg</i>	+	<i>arg</i>	<i>arg</i>	<i>arg</i>	<i>arg</i>
<i>arg</i>	+	<i>arg</i>	<i>arg</i>	<i>arg</i>	<i>arg</i>

- 35.** Suponha que você fez um cruzamento em *Neurospora* usando um mutante que tem três sítios mutantes, chamados 1, 2 e 3, no mesmo gene que estão igualmente espaçados pelo gene de 2 kb.

1 2 3 × +++

Explique a provável origem dos seguintes dois ascos:

1 2 3	1 2 3
1 2 3	1 2 3
+++	1 2 +
+++	1 2 +
+++	+++
+++	+++
+++	+++
+++	+++

36. O modelo de quebra bifilamentar ilustrado neste capítulo gera um heterodúplice. Outros modelos geram dois heterodúpliques idênticos na mesma meiose; logo, as cromátides são parentais 1, heterodúpli-

ce, heterodúplice, parental 2. Que padrões de óctade seriam produzidos das várias combinações de reparo e não-reparo desses dois heterodúpliques de mal pareamento?

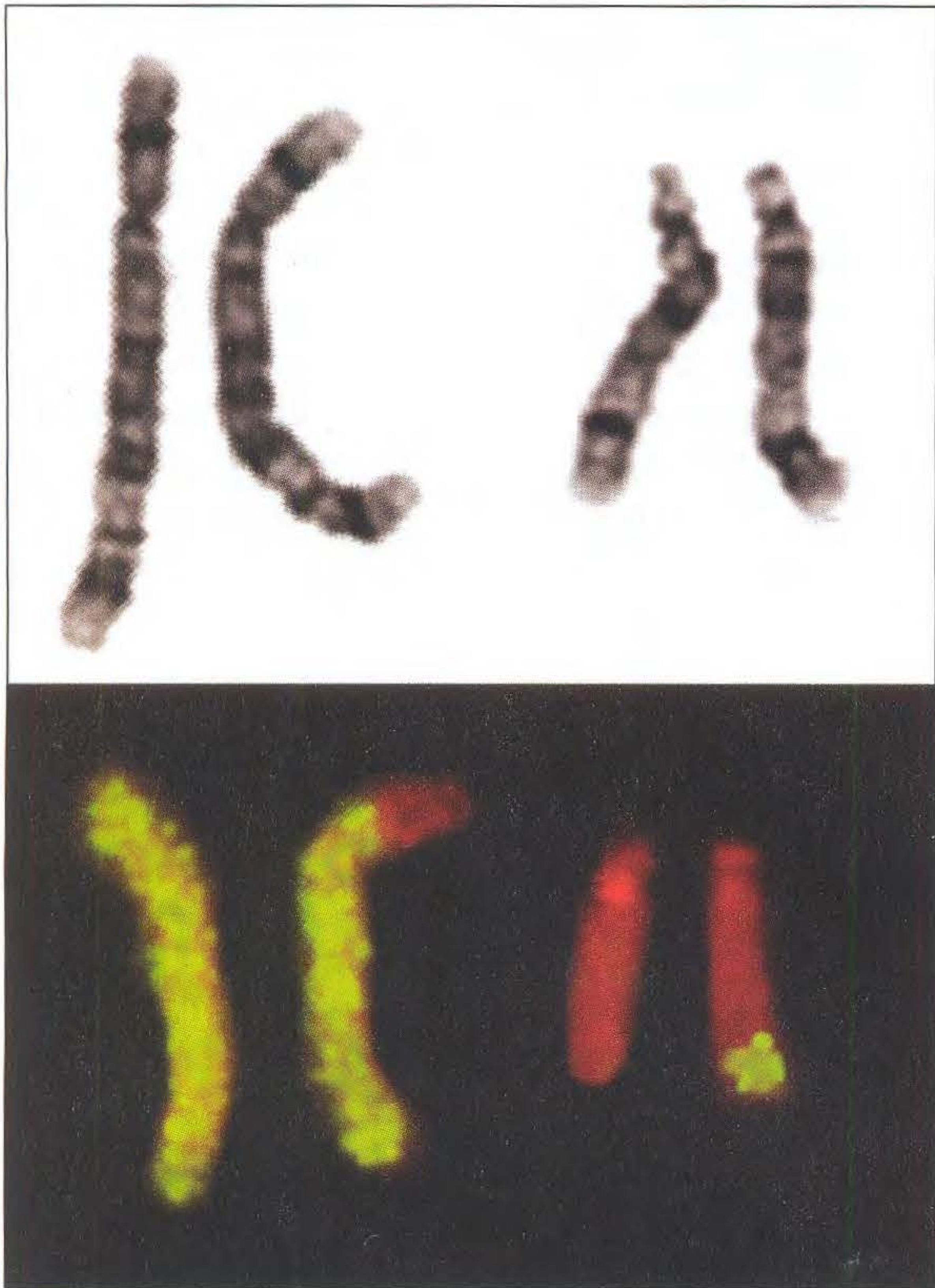
EXPLORANDO GENOMAS Um Tutorial de Bioinformática Baseado na Web

Explorando o Projeto de Anatomia do Câncer

O projeto de Cancer Anatomy do NCBI (National Center for Biotechnology Information) é um banco de dados especializado em focar um subgrupo particular de dados. Ele une, em formato gráfico e pesquisável, todas as informações sobre

genes individuais, posições de mapa e locais cromossômicos para genes associados ao câncer. No tutorial Genomics em www.whfreeman.com/iga9e, exploramos os dados de expressão gênica, mutações gênicas e anomalias cromossômicas associadas a vários tipos de tumor.

Alterações Cromossômicas em Larga Escala



Uma translocação recíproca demonstrada por coloração de cromossomos. Uma suspensão de cromossomos de muitas células é passada por um aparelho eletrônico que distribui os cromossomos pelo tamanho. O DNA é extraído de cromossomos individuais, desnaturado, ligado a um dos vários corantes fluorescentes e, então, adicionado a cromossomos parcialmente desnaturados em uma lâmina. O DNA fluorescente “encontra” seu próprio cromossomo e liga-se ao longo de seu tamanho por complementariedade de bases, “pintando-o”. Nesse exemplo, um corante verde e um vermelho são usados para pintar cromossomos diferentes. A figura mostra preparações não-coloridas e coloridas (abaixo). A preparação pintada mostra um cromossomo verde normal, um vermelho normal e dois que trocaram segmentos. [Addenbrookes Hospital/Photo Researchers.]

Um casal jovem está planejando ter filhos. O marido sabe que sua avó teve um filho com síndrome de Down no segundo casamento. A síndrome de Down é um conjunto de distúrbios físicos e mentais causados pela presença de um cromossomo 21 (Figura 16.1). Não há disponibilidade de registros, no século XX, que tenham ocorrido, mas o casal não conhece outros casos de Down em suas famílias.

O casal ouviu que a síndrome de Down resulta de uma rara chance de erro na produção de ovócitos, e, portanto, concluíram que têm uma chance pequena de ter um filho afetado. Eles decidiram ter um filho. Seu primeiro

16

Perguntas Fundamentais

- Como os poliplóides ($3n$, $4n$ e assim por diante) surgem, e quais são suas propriedades?
- Como surgem os aneuplóides ($2n - 1$, $2n + 1$ e assim por diante), e quais são suas propriedades?
- Como surgem as duplicações e deleções, e quais são suas propriedades?
- Como surgem as inversões e translocações, e quais são suas propriedades?
- Qual é a relevância de tais mudanças para os seres humanos?

Tópicos

- 16.1 Mudanças no número de cromossomos
- 16.2 Mudanças na estrutura de cromossomos
- 16.3 Incidência geral de mutações cromossômicas humanas

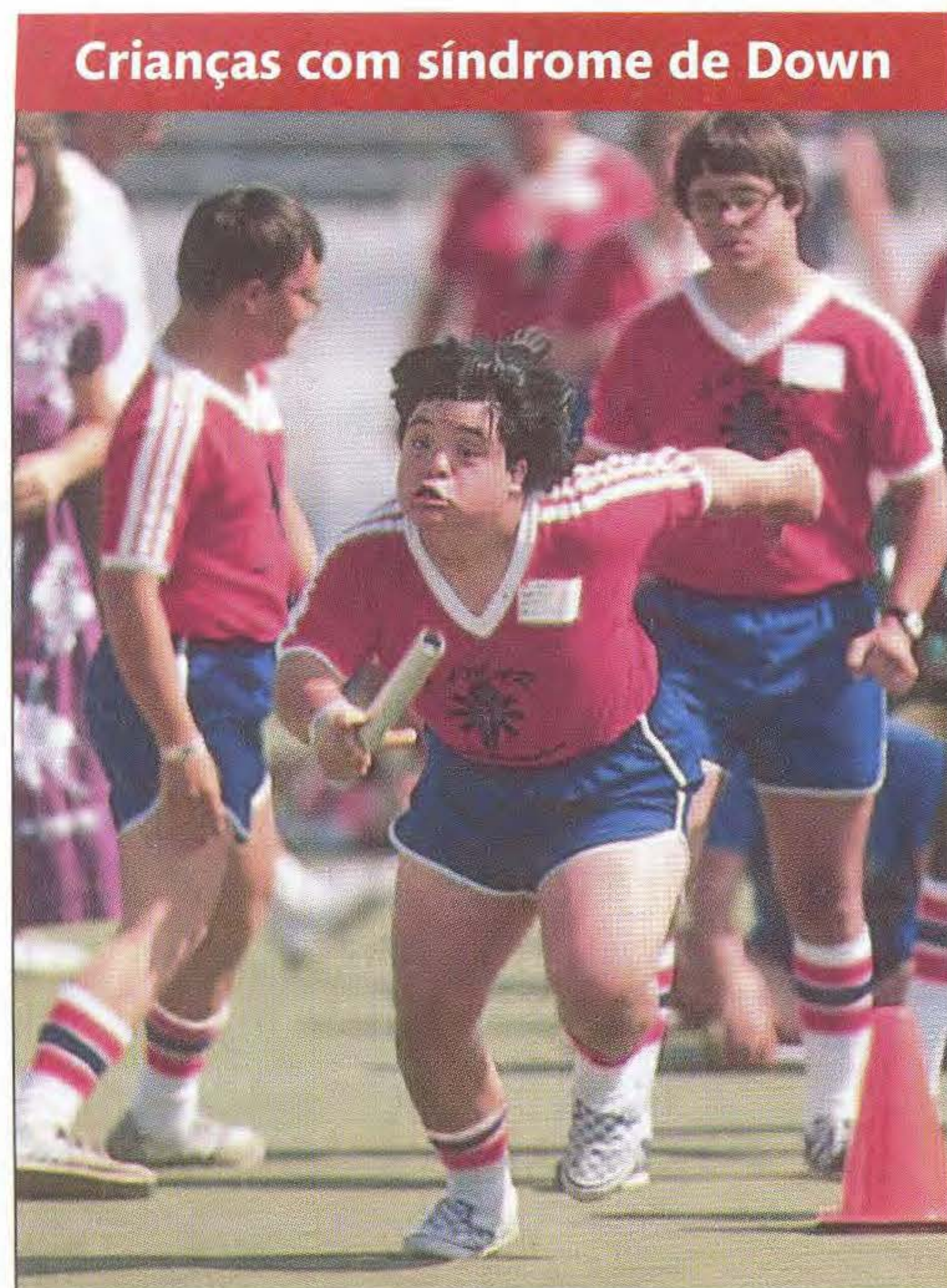


FIG. 16.1 A síndrome de Down resulta de uma cópia extra do cromossomo 21. [Bob Daemmrich/The Image Works.]

filho não é afetado, mas a concepção seguinte foi um aborto espontâneo, e o segundo filho nasceu com síndrome de Down. O fato de terem um filho com síndrome de Down foi uma coincidência ou a conexão levou a ambos terem filhos com Down? O aborto espontâneo foi significativo? Que testes podem ser necessários para investigar essa situação? A análise de tais questões é o tópico deste capítulo.

Vimos neste livro que as mutações gênicas são uma fonte importante de mudança na sequência genômica. Entretanto, o genoma também pode ser remodelado em uma escala maior por alterações na estrutura cromossômica ou por mudanças no número de cópias dos cromossomos em uma célula. Essas variações em larga escala são chamadas de **mutações cromossômicas** para distinguir-se das mutações gênicas. De modo amplo, as mutações gênicas são definidas como mudanças que ocorrem dentro de um gene, enquanto as mutações cromossômicas são mudanças em uma região cromossômica que envolvem vários genes. As mutações gênicas nunca são detectáveis microscopicamente; um cromossomo tendo uma mutação gênica parece igual, ao microscópio, a outro que tenha uma cópia normal do alelo. Em contraste, muitas mutações cromossômicas são detectáveis com o uso de microscopia. As mutações cromossômicas podem ser detectadas por microscopia, por análise genética ou molecular, ou por uma combinação de todas as técnicas. As mutações cromossômicas foram mais bem caracterizadas em eucariontes, e todos os exemplos neste capítulo são desse grupo.

As mutações cromossômicas são importantes por vários aspectos biológicos. Primeiro, elas podem ser fontes de *insight* de como os genes atuam em conjunto em uma escala genômica. Segundo, elas revelam várias características importantes da meiose e arquitetura cromossômica. Terceiro, elas

constituem ferramentas úteis para manipulação genômica experimental. Quarto, elas são fontes de *insight* em processos evolutivos. Quinto, as mutações cromossômicas são encontradas regularmente em humanos, e algumas dessas mutações causam doenças genéticas.

Muitas mutações cromossômicas causam anomalias na célula e funcionamento do organismo. Em sua maioria, essas anomalias são baseadas em mudanças no *número de genes* ou *posição gênica*. Em alguns casos, uma mutação cromossômica resulta de quebra cromossômica. Se a quebra ocorre dentro de um gene, o resultado é a *perturbação* funcional desse gene.

Para nossas finalidades, devemos dividir as mutações cromossômicas em dois grupos: mudanças no *número* de cromossomos e mudanças na *estrutura* do cromossomo. Esses dois grupos representam dois tipos fundamentalmente diferentes de eventos. As mudanças no número de cromossomos não estão associadas a alterações estruturais de nenhuma das moléculas de DNA da célula. De preferência, o *número* dessas moléculas de DNA que é alterado, e essa mudança de número é a base de seus efeitos genéticos. Mudanças na estrutura do cromossomo, por outro lado, resultam em novos arranjos de sequência dentro de uma ou mais duplas hélices de DNA. Esses dois tipos de mutações cromossômicas são ilustrados na Figura 16.2, que é um resumo dos tópicos deste capítulo. Começaremos explorando a natureza e as consequências das mudanças no número de cromossomos.

► 16.1 Mudanças no Número de Cromossomos

Em genética como um todo, alguns tópicos influenciam as questões humanas tão diretamente quanto as mudanças no número de cromossomos presentes em nossas células. Há também o fato de que um grupo de distúrbios genéticos comuns resulta da presença de um número anormal de cromossomos. Embora esse grupo de distúrbios seja pequeno, ele contribui para uma grande proporção de problemas de saúde geneticamente determinados que afligem os humanos. É também relevante para os humanos o papel das mutações cromossômicas no cruzamento de plantas: os que cultivam plantas rotineiramente manipulam o número de cromossomos para melhorar cultivos agrícolas comercialmente importantes.

As mudanças no número de cromossomos são de dois tipos básicos: mudanças em conjuntos cromossômicos *inteiros*, resultando em uma condição chamada de *euploidia aberrante*, e mudanças em *partes* de conjuntos cromossômicos, resultando em uma condição chamada de *aneuploidia*.

Euploidia aberrante

Os organismos com múltiplos do conjunto cromossômico básico (genoma) são chamados de **euplóides**. Você aprendeu, nos capítulos anteriores, que os eucariontes familiares, tais como plantas, animais e fungos, levam em suas células ou um conjunto cromossômico (haploidia), ou dois conjuntos cromossômicos (diploidia). Nessas espécies, tanto o estado

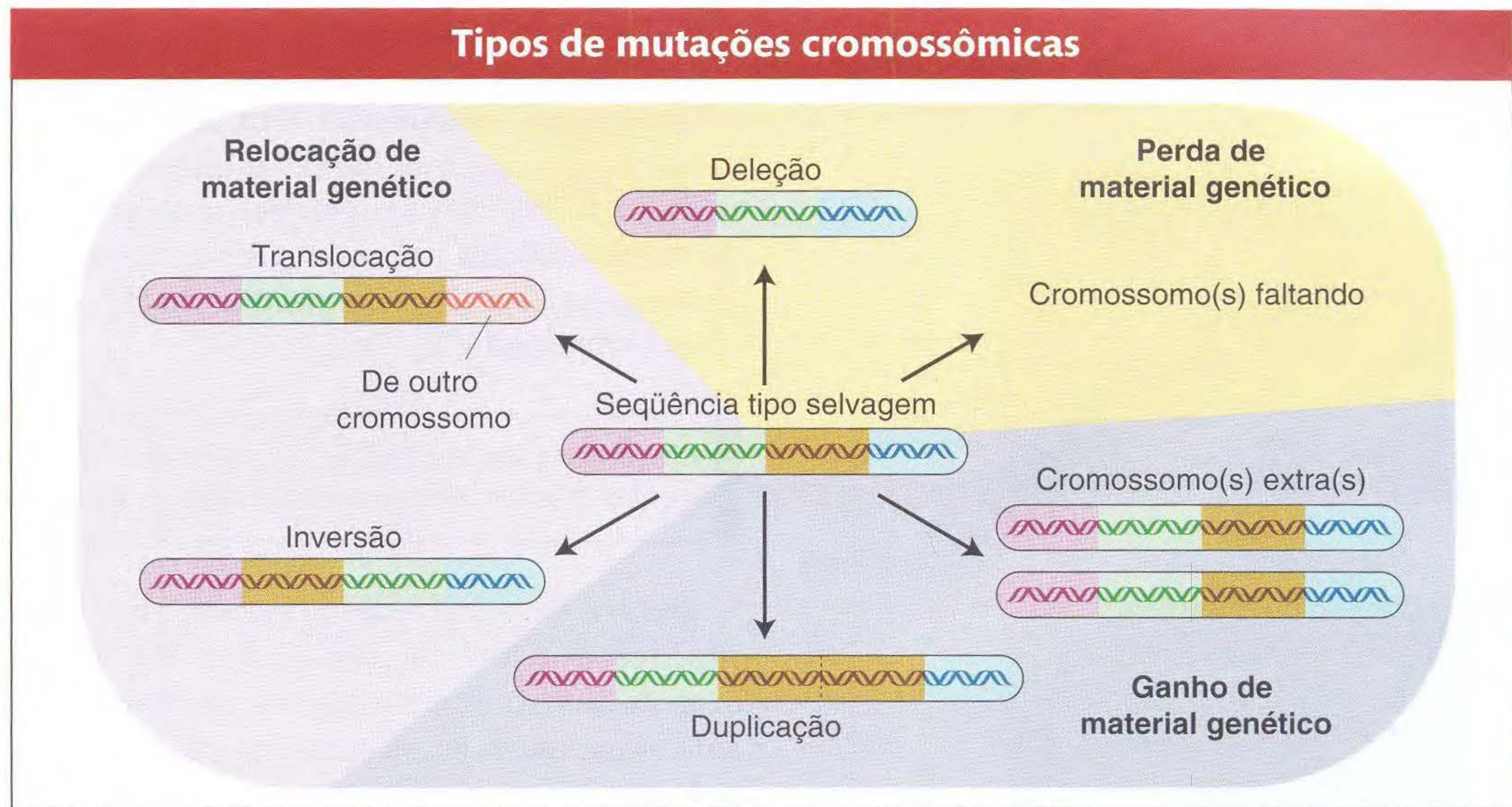


FIG. 16.2 A ilustração é dividida em três regiões coloridas para mostrar os principais tipos de mutações cromossômicas que podem ocorrer: a perda, ganho ou relocação de cromossomos inteiros ou segmentos cromossômicos. O cromossomo tipo selvagem é mostrado no centro.

haplóide quanto o diplóide são casos de euploidia normal. Os organismos que têm mais ou menos o número normal de conjuntos são euplóides aberrantes. Os **poliplóides** são organismos que têm mais de dois conjuntos cromossômicos. Eles podem ser representados por $3n$ (**triplóides**), $4n$ (**tetraplóides**), $5n$ (**pentaplóides**), $6n$ (**hexaplóides**), e assim por diante. (O número de conjuntos cromossômicos é chamado de ploidia ou nível de ploidia.) Um membro individual de uma espécie normalmente diplóide que tem só um conjunto cromossômico (n) é chamado de **monoplóide** para distinguir-se de um membro individual de uma espécie normalmente haplóide (também n). Os exemplos dessas condições são mostrados nas primeiras quatro fileiras do Quadro 16.1.

Monoplóides Os machos de abelhas, vespas e formigas são monoplóides. Nos ciclos de vida normais desses insetos, os machos desenvolvem-se por **partenogênese** (o desenvolvimento de um tipo especializado de ovócito não-fertilizado em um embrião sem a necessidade de fertilização). Na maioria das outras espécies, entretanto, os zigotos monoplóides não se desenvolvem. O motivo é que praticamente todos os membros de uma espécie diplóide levam várias mutações deletérias recessivas, em conjunto chamadas de **“carga genética”**. Os alelos deletérios recessivos são mascarados pelos alelos tipo selvagem na condição diplóide, mas são automaticamente expressos em um monoplóide derivado de um diplóide. Os monoplóides que se desenvolvem em estágios avançados são anormais. Se eles sobrevivem até a vida adulta, suas célu-

Quadro 16.1 Constituições Cromossômicas em um Organismo Normalmente Diplóide com Três Cromossomos (Identificados como A, B e C) no Conjunto Básico

Nome	Designação	Constituição	Número de cromossomos
<i>Euplóides</i>			
Monoplóide	n	A B C	3
Diplóide	$2n$	AA BB CC	6
Triplóide	$3n$	AAA BBB CCC	9
Tetraplóide	$4n$	AAAA BBBB CCCC	12
<i>Aneuplóides</i>			
Monossômico	$2n - 1$	A BB CC	5
		AA B CC	5
		AA BB C	5
		AAA BB CC	7
		AA BBB CC	7
		AA BB CCC	7
Trissômico	$2n + 1$		

Uma maior ploidia produz tamanho maior

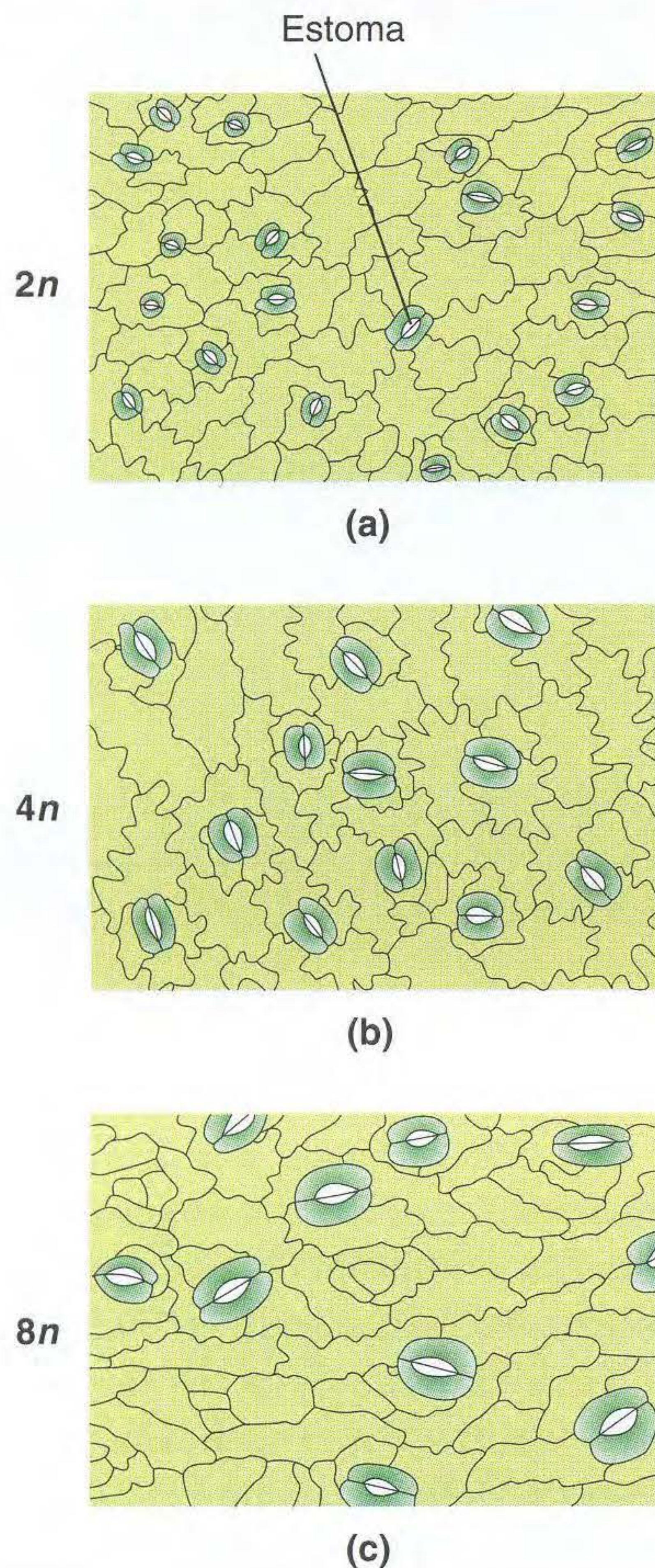


FIG. 16.3 Células da epiderme de folhas de plantas de tabaco com aumento de ploidia. O tamanho da célula aumenta, particularmente o tamanho dos estomas, com o aumento de ploidia. (a) Diplóide; (b) tetraplóide; (c) octoplóide. [De W. Williams, *Genetic Principles and Plant Breeding*. Blackwell Scientific Publications, Ltd., 1964.]

las germinativas não podem fazer normalmente uma meiose, pois os cromossomos não têm parceiros de pareamento. Assim, os monoplóides são caracteristicamente estéreis. (Machos de abelhas, vespas e formigas contornam a meiose; nesses grupos, os gametas são produzidos por *mitose*.)

Poliplóides A poliploidia é muito comum em plantas, mas raramente em animais (por motivos que consideraremos depois). De fato, um aumento no número de conjuntos cromossômicos tem sido um fator importante na origem de novas espécies de plantas. A evidência para esse benefício é mostrada na Figura 19.16, que apresenta a frequência de distribuição do número haplóide (o número no conjunto básico de cromossomos) de espécies de plantas dicotiledôneas. Acima do número haplóide de cerca de 12, números ímpares são muito mais comuns que números pares. Esse padrão é uma consequência da origem poliplóide de muitas espécies de plantas, pois a duplicação e a reduplicação de um número podem dar origem a números pares. As espécies animais não apresentam tal distribuição devido à relativa raridade de animais poliplóides.

Em euplóides aberrantes, sempre há uma correlação entre o número de cópias do conjunto cromossômico e o tamanho do organismo. Um organismo tetraplóide, por exemplo, tipicamente parece muito similar à sua contraparte diplóide em suas proporções, porém o número tetraplóide é maior, como um todo e em suas partes componentes. Quanto maior o nível de ploidia, maior o tamanho do organismo (Figura 16.3).

Os poliplóides são, em geral, maiores e têm partes componentes maiores do que seus correlatos diplóides.

Dentre os poliplóides, devemos distinguir os **autopoliplóides**, que têm conjuntos cromossômicos múltiplos com origem dentro de uma espécie, e **alopoliplóides**, que têm conjuntos de duas ou mais espécies diferentes. Os alopólóides formam-se apenas entre espécies proximamente relacionadas. Entretanto, os conjuntos cromossômicos diferentes são apenas **homeólogos** (parcialmente homólogos), não totalmente homólogos como são nos autopoliplóides.

Autopoliplóides Os triplóides ($3n$) são geralmente autopoliplóides. Eles surgem espontaneamente na natureza, mas podem ser construídos por geneticistas do cruzamento de um $4n$ (tetraplóide) e um $2n$ (diplóide). Os gametas $2n$ e n produzidos pelo tetraplóide e diplóide, respectivamente, unem-se para formar um triplóide $3n$. Os triplóides são caracteristicamente estéreis. O problema (que também é verdadeiro nos monoplóides) está na presença de cromossomos não-pareados na meiose. Os mecanismos moleculares para sinapse, ou verdadeiro pareamento, determinam que, em um triplóide, o pareamento pode ocorrer apenas entre dois dos três cromossomos de cada tipo (Figura 16.4). Os homólogos pareados (**bivalentes**) segregam-se para pólos opostos, mas os homólogos não-pareados (**univalentes**) passam para um dos pólos aleatoriamente. Em um **trivalente**, um grupo pareado de três, os centrômeros pareados segregam-se como um bivalente, e o não-pareado como um univalente. Essas segregações ocorrem para cada trinca de cromossomos; assim, para qualquer tipo cromossômico, o gameta pode receber um ou dois cromossomos. É improvável que um gameta receba *dois* de *cada* tipo cromossômico ou que receba *um* de *cada* tipo cromossômico. Assim, a probabilidade é que os gametas tenham números cromossômicos intermediários entre o haplóide e o número diplóide. Tais genomas são de um tipo chamado de **aneuplóide** (“não-euplóide”).

Os gametas aneuplóides geralmente não dão origem a prole viável. Nas plantas, os grãos de pólen aneuplóides são geralmente inviáveis e, assim, incapazes de fertilizar o gameta feminino. Em qualquer organismo, os zigotos que podem surgir da fusão de um gameta haplóide e um aneuplóide serão eles mesmos aneuplóides, e tipicamente esses zigotos também são inviáveis. Examinaremos o motivo subjacente para a inviabilidade dos aneuplóides quando considerarmos o balanço mais adiante neste capítulo.

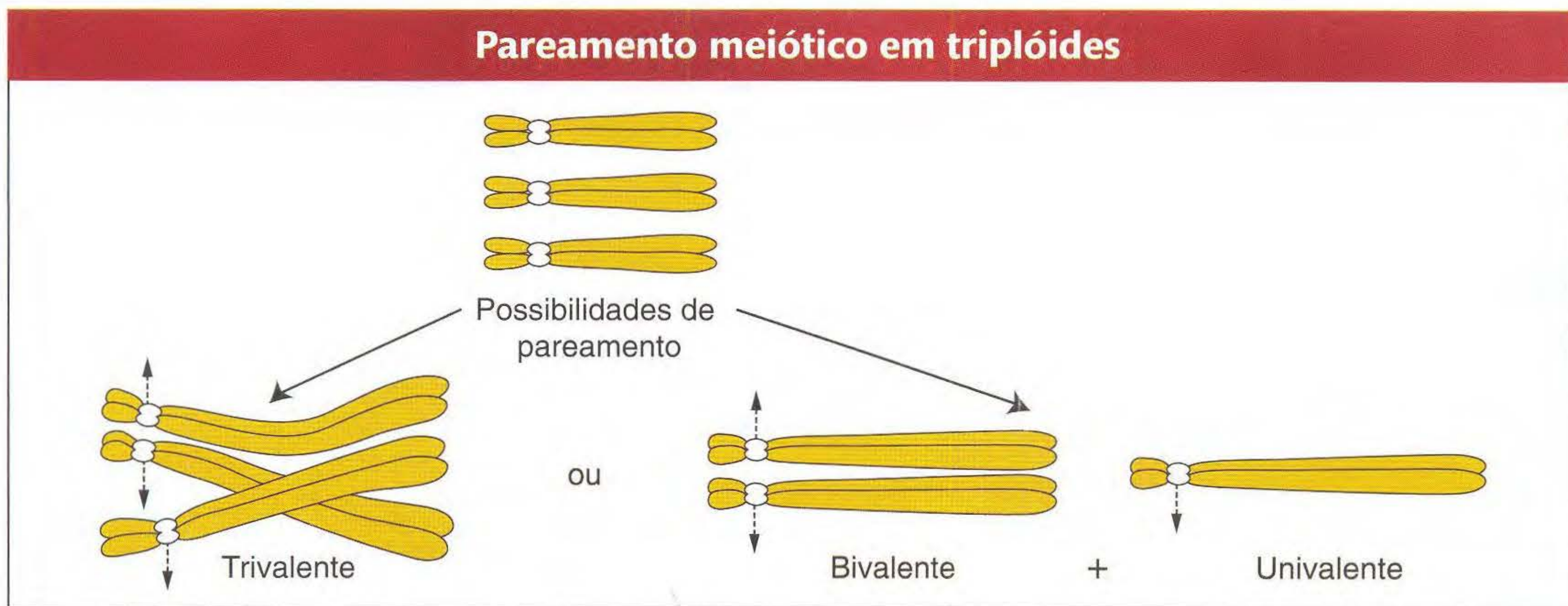


FIG. 16.4 Os três cromossomos homólogos de um triplóide podem parear de dois modos na meiose: como um trivalente ou como um bivalente mais um univalente.

Os poliplóides com números ímpares de conjuntos cromossômicos, tais como triplóides, são estéreis ou altamente inférteis porque seus gametas e prole são aneuplóides.

Os autotetraplóides surgem pela duplicação do complemento $2n$ para $4n$. Essa duplicação pode ocorrer espontaneamente, mas também pode ser induzida artificialmente aplicando agentes químicos que perturbam a polimerização de microtúbulos. Como dito no Capítulo 2, a segregação cromossômica é ativada pelas fibras do fuso, que são polímeros da proteína tubulina. Assim, a perturbação da polimerização dos microtúbulos bloqueia a segregação cromossômica. O tratamento químico é normalmente aplicado ao tecido somático durante a formação das fibras do fuso nas células que sofrem divisão. O tecido poliplóide resultante (tal como um ramo poliplóide de uma planta) pode ser detectado examinando-se cromossomos corados do tecido ao microscópio. Tal ramificação pode ser removida e usada como uma muda para gerar uma planta poliplóide ou deixada produzir flores, que, quando autofecundadas, produziriam prole poliplóide. Um agente antitubulina comumente usado é a colchicina, um alcalóide extraído de açafrão. Nas células tratadas com colchicina, ocorre a fase S do ciclo celular, mas a segrega-

ção cromossômica ou divisão celular não. À medida que a célula tratada entra na telófase, forma-se uma membrana nuclear ao redor de todo o conjunto duplo de cromossomos. Assim, tratar as células diplóides ($2n$) com colchicina por um ciclo celular leva a tetraplóides ($4n$) com, exatamente, quatro cópias de cada tipo de cromossomo (Figura 16.5). O tratamento para um ciclo celular adicional produz octoplóides ($8n$), e assim por diante. Esse método funciona tanto em células vegetais quanto animais, mas, geralmente, as plantas parecem ser muito mais tolerantes à poliploidia. Note que todos os alelos no genótipo estão duplicados. Portanto, se uma célula diplóide do genótipo $A/a ; B/b$ é duplicada, o autotetraplóide resultante será do genótipo $A/A/a/a ; B/B/b/b$.

Como quatro é um número par, os autotetraplóides podem ter uma meiose regular, embora esse resultado nem sempre seja o caso. O fator crucial é como quatro cromossomos de cada conjunto formam par e se segregam. Existem várias possibilidades, como mostrado na Figura 16.6. Se os cromossomos formam pares como bivalentes ou quadrivalentes, os cromossomos segregam-se normalmente, produzindo gametas diplóides. A fusão de gametas na fertilização regenera o estado tetraplóide. Se forem formados trivalentes, a segregação leva a gametas aneuplóides não-funcionais e, portanto, à esterilidade.

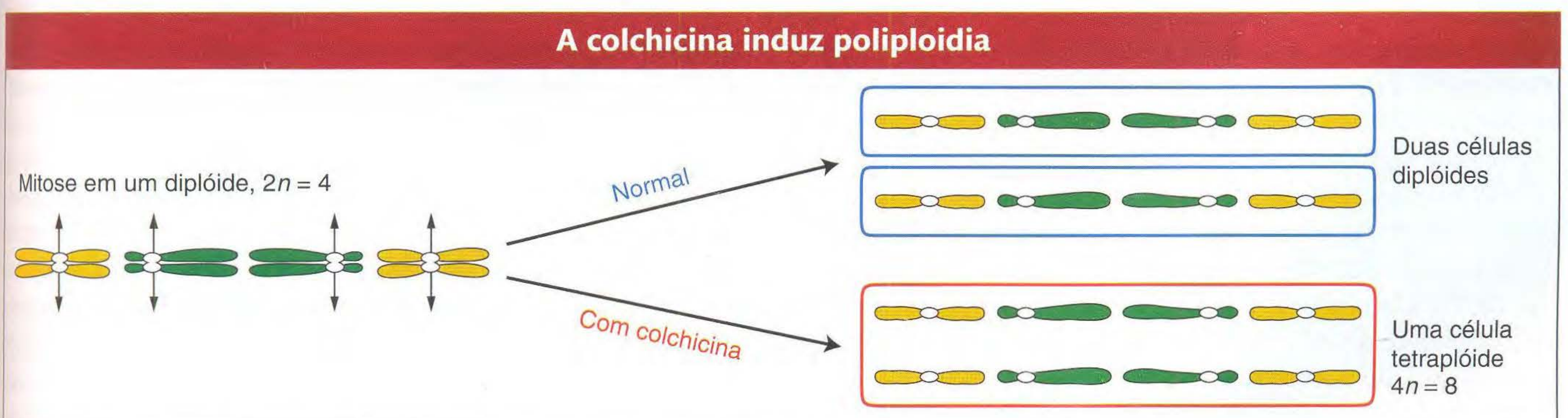


FIG. 16.5 A colchicina pode ser aplicada para gerar um tetraplóide a partir de um diplóide. A colchicina adicionada a células mitóticas durante a metáfase e anáfase perturba a formação das fibras do fuso, impedindo a migração de cromátides após o centrômero ter se dividido. É criada uma única célula que contém pares de cromossomos idênticos, homozigotos em todos os loci.

Pareamento cromossômico em um autotetraplóide

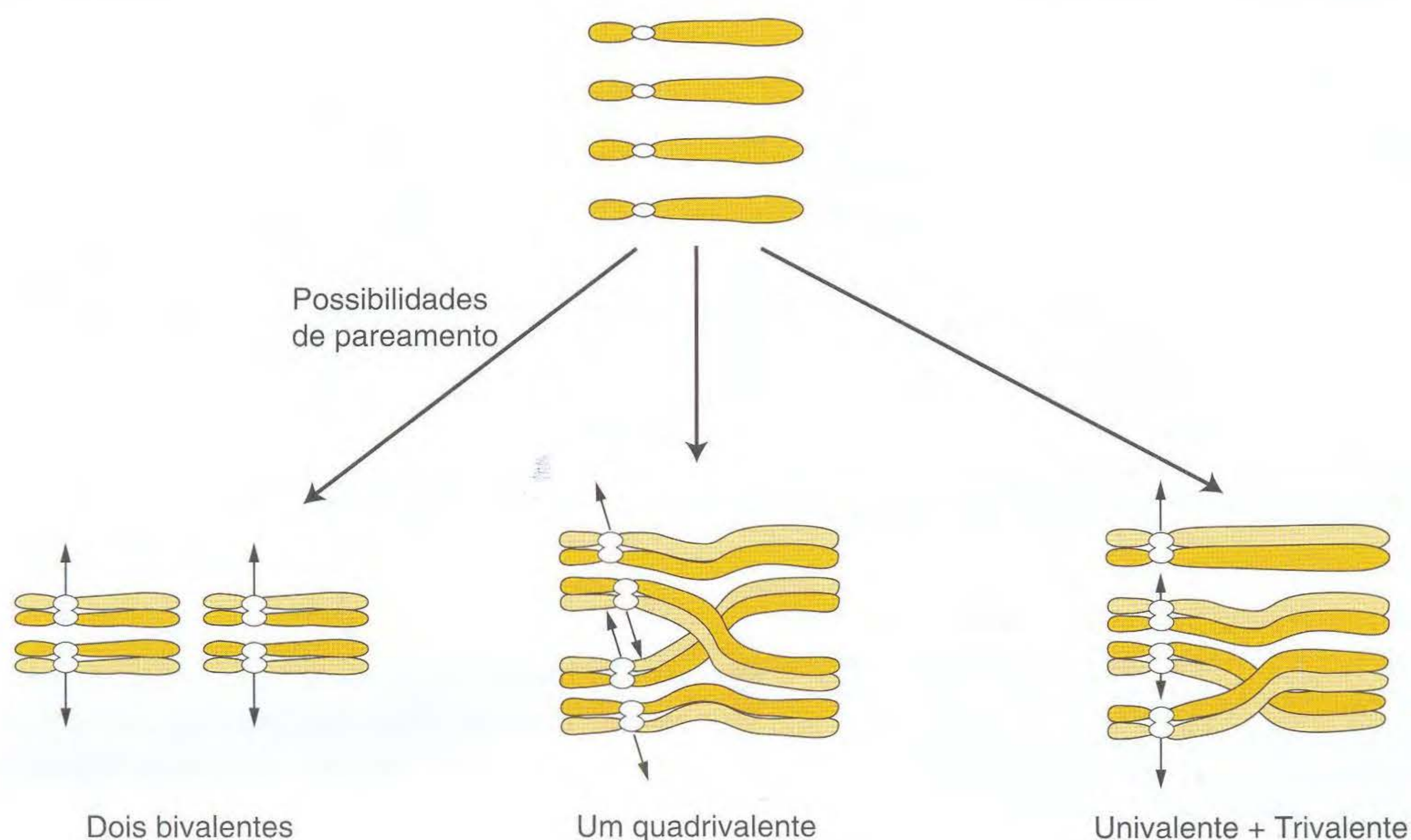


FIG. 16.6 Existem três possibilidades diferentes de pareamento na meiose nos tetraplóides. Os quatro cromossomos homólogos podem parear como dois bivalentes ou como um quadrivalente, e cada um pode produzir gametas funcionais. Uma terceira possibilidade, um trivalente mais um univalente, produz gametas não-funcionais.

Que proporções genéticas são produzidas por um autotetraplóide? Suponha, por simplificação, que o tetraplóide forme apenas bivalentes. Se começarmos com uma planta tetraplóide $A/A/a/a$ e a autofecundarmos, que proporção da prole será $a/a/a/a$? Primeiro precisamos deduzir a frequência de gametas a/a porque esse tipo é o único que pode produzir um homozigoto recessivo. Os gametas a/a podem surgir apenas se ambos os pareamentos forem A com a , e, então, ambos os alelos a devem segregar para o mesmo pólo. Vamos usar o seguinte experimento teórico para calcular as frequências dos resultados possíveis. Considere as opções do ponto de vista de um dos cromossomos a face às opções de pareamento com o outro cromossomo a ou com um dos dois cromossomos A . Se o pareamento for aleatório, há dois terços de chance de que ele forme par com um cromossomo A . Se o fizer, então o pareamento dos dois cromossomos restantes necessariamente será de A com a porque são os únicos cromossomos restantes. Com esses dois pareamentos A com a existem duas segregações igualmente possíveis, e, no geral, um quarto dos produtos conterá ambos os alelos a em um pólo. Assim, a probabilidade de um gameta a/a será $2/3 \times 1/4 = 1/6$. Assim, se os gametas parearem aleatoriamente, a probabilidade de um zigoto $a/a/a/a$ será $1/6 \times 1/6 = 1/36$ e, por subtração, a probabilidade de $A/-/-/-$ será $35/36$. Portanto, é esperada uma proporção fenotípica de 35:1.

Alopoliplóides Um alopoliplóide é uma planta que é híbrida de duas ou mais espécies, contendo duas ou mais cópias de cada um dos genomas. O alopoliplóide prototípico foi um alo-

tetraplóide sintetizado por Georgi Karpechenko em 1928. Ele queria fazer um híbrido fértil que tivesse as folhas de couve (*Brassica*) e as raízes de rabanete (*Raphanus*), porque eram partes agriculturalmente importantes de cada planta. Cada uma dessas duas espécies tinha 18 cromossomos, e assim $2n_1 = 2n_2 = 18$ e $n_1 = n_2 = 9$. As espécies são proximamente correlacionadas, o suficiente para permitir o entrecruzamento. A fusão de um gameta n_1 e um gameta n_2 produziu uma prole híbrida viável de indivíduos com a constituição $n_1 + n_2 = 18$. Entretanto, esse híbrido era funcionalmente estéril porque os 9 cromossomos do genitor couve eram suficientemente diferentes dos cromossomos de rabanete de modo a não haver sinapse e segregação normal na meiose, e, portanto, o híbrido não podia produzir gametas funcionais.

Ocasionalmente, uma parte da planta híbrida produziu algumas sementes. Plantadas, essas sementes produziram indivíduos férteis com 36 cromossomos. Todos esses indivíduos eram alopoliplóides. Aparentemente, tinham sido derivados de duplicação cromossômica espontânea acidental, passando para $2n_1 + 2n_2$ em uma região do híbrido estéril, supostamente no tecido que se tornou uma flor e sofreu meiose para produzir gametas. No tecido $2n_1 + 2n_2$, há um parceiro de pareamento para cada cromossomo, e foram produzidos gametas funcionais do tipo $n_1 + n_2$. Esses gametas se fundem para dar uma prole alopoliplóide $2n_1 + 2n_2$, que também era fértil. Esse tipo de alopoliploidia às vezes é chamado de **anfidiplóide**, ou diploidia duplicada (Figura 16.7). Tratando o híbrido estéril com colchicina, aumentam muito as chances de que os conjuntos cromossômicos se dupliquem.

Origem de *Raphanobrassica* anfidiplóide

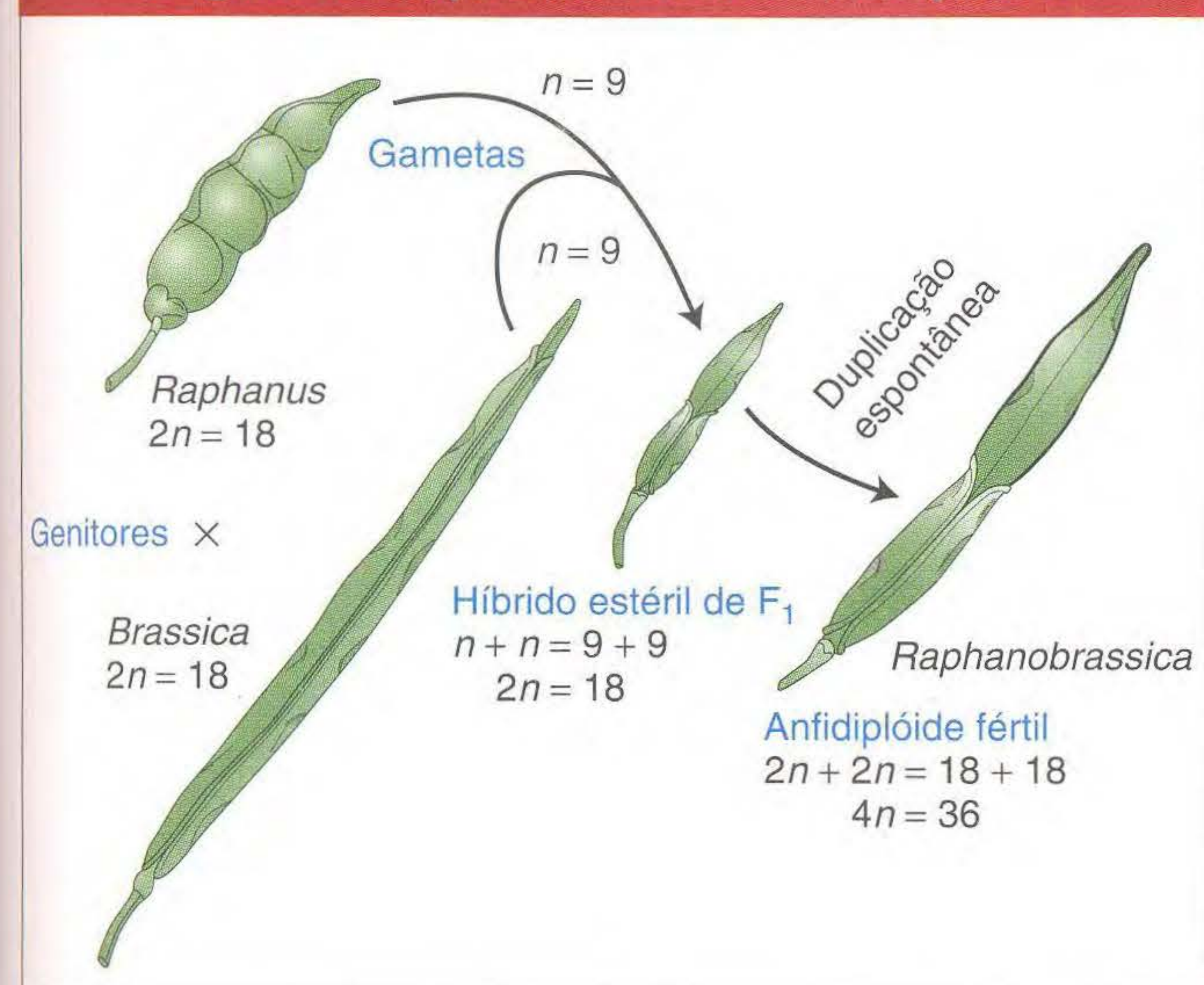


FIG. 16.7 Na prole de um cruzamento de couve (*Brassica*) e rabanete (*Raphanus*), o anfidiplóide fértil surgiu da duplicação espontânea no híbrido estéril $2n=18$. [De A.M. Srb, R.D. Owen, and R.S. Edgar, *General Genetics*. 2nd ed. Copyright 1965 by W.H. Freeman and Company. Adaptado de G. Karpechenko, *Z. Indukt. Abst. Vererb.* 48, 1928, 27.]

Os anfidiplóides hoje são sintetizados rotineiramente desse modo. (Infelizmente para Karpechenko, seu anfidiplóide tinha raízes de couve e folhas de rabanete.)

Quando o alopoliplóide de Karpechenko foi cruzado com uma das espécies parentais, couve ou rabanete, resultou uma prole estéril. A prole do cruzamento com couve era $2n_1 + n_2$, constituída de gametas $n_1 + n_2$ gametas do alopoliplóide e um gameta n_1 de couve. Os cromossomos n_2 não tinham parceiros de pareamento; assim, não podia ocorrer uma meiose normal, e a prole era estéril. Portanto, Karpechenko havia efetivamente criado uma nova espécie, sem possibilidade de troca de genes com couve ou com rabanete. Ele chamou essa nova planta de *Raphanobrassica*.

Na natureza, a alopoliploidia parece ter sido uma força importante na evolução de novas espécies de plantas. Um exemplo convincente é mostrado pelo gênero *Brassica*, como ilustrado na Figura 16.8. Aqui, três espécies diferentes tinham se hibridizado em todas as possíveis combinações de pares para formar novas espécies anfidiplóides. A poliploidia natural foi vista como uma ocorrência rara, mas trabalhos recentes mostraram que ela é um evento recorrente em muitas espécies de plantas. O uso de marcadores de DNA possibilitou mostrar que os poliplóides, em qualquer população ou área que pareçam ser a mesma, são o resultado de muitas fusões passadas independentes entre indivíduos geneticamente distintos das mesmas duas espécies parentais. Uma estimativa de 50% de todas as plantas angiospermas são poliplóides, resultando de auto ou alopoliploidia. Como resultado de múltiplas poliploidizações, a quantidade de variação alélica dentro de uma espécie poliplóide é muito maior do que se imaginava, talvez contribuindo para seu potencial de adaptação.

Um alopoliplóide natural particularmente interessante é o trigo do pão, *Triticum aestivum* ($6n=42$). Estudando seus parentes selvagens, os geneticistas reconstruíram uma provável história evolutiva dessa planta. A Figura 16.9 mostra que o trigo do pão é composto de dois conjuntos de três genomas ancestrais. Na meiose, o pareamento é sempre entre os homólogos do mesmo genoma ancestral. Assim, na meiose do trigo do pão, existem sempre 21 bivalentes.

As plantas alopoliplóides também podem ser produzidas artificialmente fundindo células diplóides de espécies diferentes. Primeiro, as paredes de duas células diplóides são removidas pelo tratamento com uma enzima, e as membranas das duas células fundem-se e tornam-se uma só. Os núcleos em geral se fundem, resultando no poliplóide. Se a célula é nutrida com os hormônios apropriados e nutrientes, ela se divide para tornar-se uma pequena plantinha alopoliplóide, que pode então ser transferida para o solo.

Plantas alopoliplóides podem ser sintetizadas cruzando espécies correlacionadas e duplicando os cromossomos do híbrido ou fundindo células diplóides.

Aplicações agrícolas As variações no número de cromossomos foram exploradas para criar novas linhagens de plantas com as características desejáveis. Seguem alguns exemplos.

Monoplóides A diploidia é uma característica inerente para os criadores de plantas. Quando eles querem induzir e selecionar novas mutações recessivas que são favoráveis para fins agrícolas, as novas mutações não podem ser detectadas a menos que sejam homólogas. Os agricultores também podem querer encontrar novas combinações favoráveis de alelos em loci diferentes, mas tais combinações favoráveis de alelos em heterozigotos serão rompidas pela recombinação na meiose. Os monoplóides fornecem um meio de contornar alguns desses problemas.

Os monoplóides podem ser artificialmente derivados de produtos da meiose nas anteras da planta. Uma célula haplóide destinada a tornar-se um grão de pólen pode ser induzida, por tratamento com frio (sujeito a temperaturas baixas), a crescer em um **embrióide**, uma pequena massa em divisão de células monoplóides. O embrióide pode ser cultivado em ágar para formar uma plântula monoplóide, que pode então ser plantada no solo e deixada amadurecer (Figura 16.10).

Os monoplóides de plantas podem ser explorados de vários modos. Em um enfoque, eles são primeiro examinados para combinações alélicas que surgiram da recombinação de alelos já presentes em um genitor diplóide heterozigoto. Assim, a partir de um genitor que é $A/a ; B/b$, pode surgir uma combinação monoplóide favorável $a ; b$. O monoplóide pode então ser submetido a duplicação cromossômica para produzir células diplóides homozigotas, $a/a ; b/b$, que são capazes de reprodução normal.

Outro enfoque é tratar células monoplóides basicamente com uma população de organismos haplóides em um procedimento de mutagênese e seleção. É isolada uma população

Origem de três espécies alopoliplóides de *Brassica*

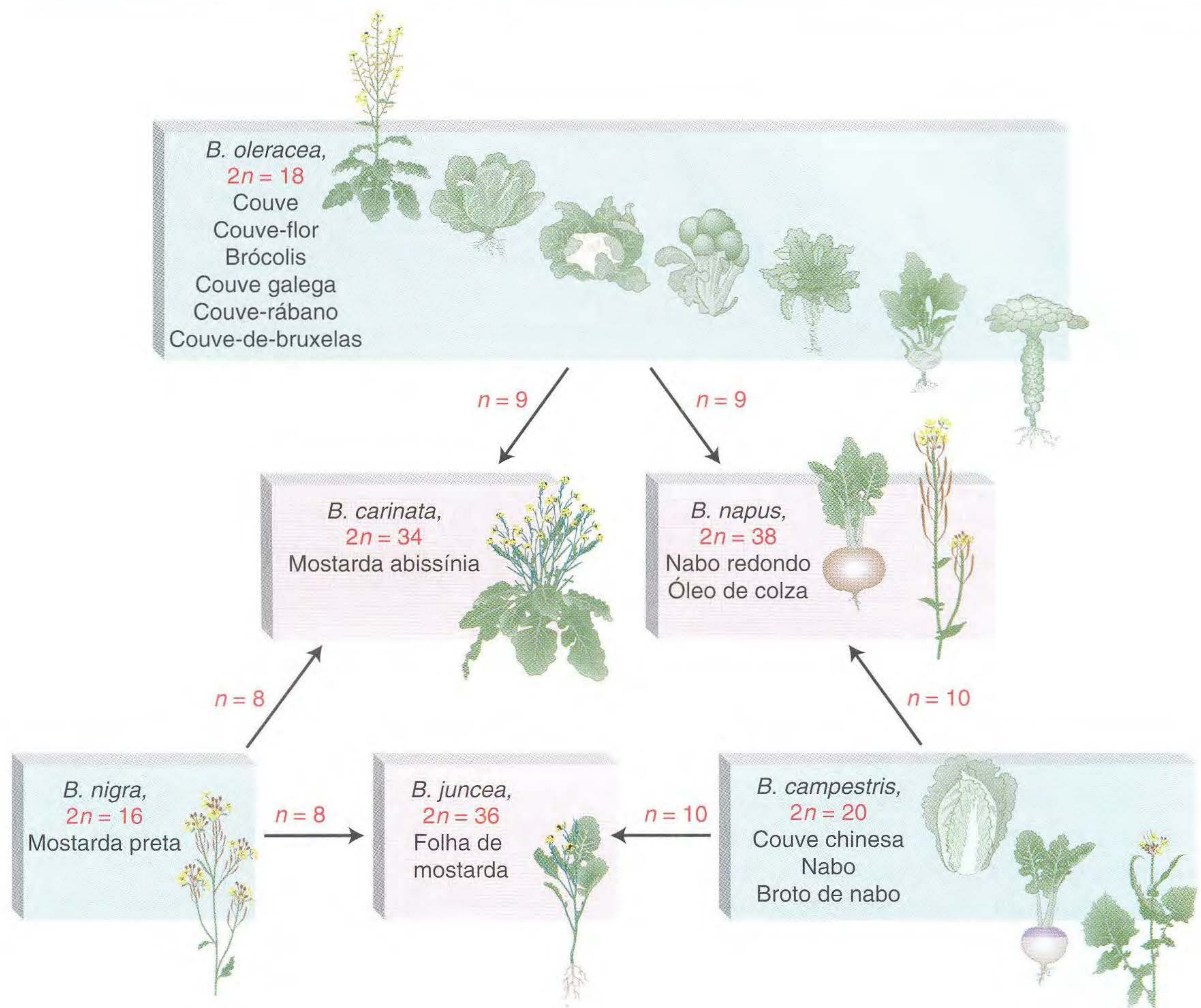


FIG. 16.8 A alopoliploidia é importante na produção de novas espécies. No exemplo mostrado, três espécies diplóides de *Brassica* (boxes azuis) foram cruzadas em diferentes combinações para produzir seus alopoliplóides (boxes rosa). Alguns dos derivados agrícolas de algumas das espécies são mostrados dentro de boxes.

de células monoplóides, suas paredes são removidas por tratamento enzimático, e elas são expostas a um mutágeno. A seguir, são plaqueadas em um meio que seleciona alguns fenótipos desejáveis. Esse enfoque foi usado para selecionar a resistência a compostos tóxicos produzidos por um parasita da planta, bem como para selecionar quanto à resistência a herbicidas usados por fazendeiros para matar pragas. As plântulas resistentes crescem em plantas monoplóides, cujo número de cromossomos pode então ser duplicado com o uso de colchicina. Esse tratamento produz tecido diplóide e, ocasionalmente, colhendo uma muda ou autofecundando uma flor, produz uma planta diplóide totalmente resistente. Essas técnicas poderosas podem evitar o processo normalmente lento de cruzamento de plantas baseado na meiose. Eles têm sido bem-sucedidamente aplicados no cultivo de plantas importantes, tais como soja e tabaco.

Os geneticistas podem criar novas linhagens de plantas produzindo monoplóides com genótipos favoráveis e, então, duplicando seus cromossomos para formar diplóides férteis, homozigotos.

Autotriplóides As bananas amplamente disponíveis no comércio são triplóides estéreis com 11 cromossomos em cada conjunto ($3n = 33$). A expressão mais óbvia da esterilidade das bananas é a ausência de sementes no fruto que comemos. (Os pontos pretos nas bananas não são sementes; as sementes das bananas são muito duras, realmente quebradoras de dentes.) As melancias sem sementes são outro exemplo de exploração comercial de triploidia em plantas.

Autotetraplóides Muitas plantas autotetraplóides foram desenvolvidas como cultivos comerciais para tirar proveito de seu aumento de tamanho (Figura 16.11). Grandes frutos e flores são particularmente favorecidos.

Alopoliplóides A alopoliploidia (formação de poliplóides entre espécies diferentes) foi importante na produção de modernos cultivos de plantas. O algodão do Novo Mundo é um alopoliplóide natural que surgiu espontaneamente, como no trigo. Os alopoliplóides também são sintetizados artificialmente para combinar as características úteis de espécies parentais em um tipo. Apenas um anfidiplóide sintético foi amplamente

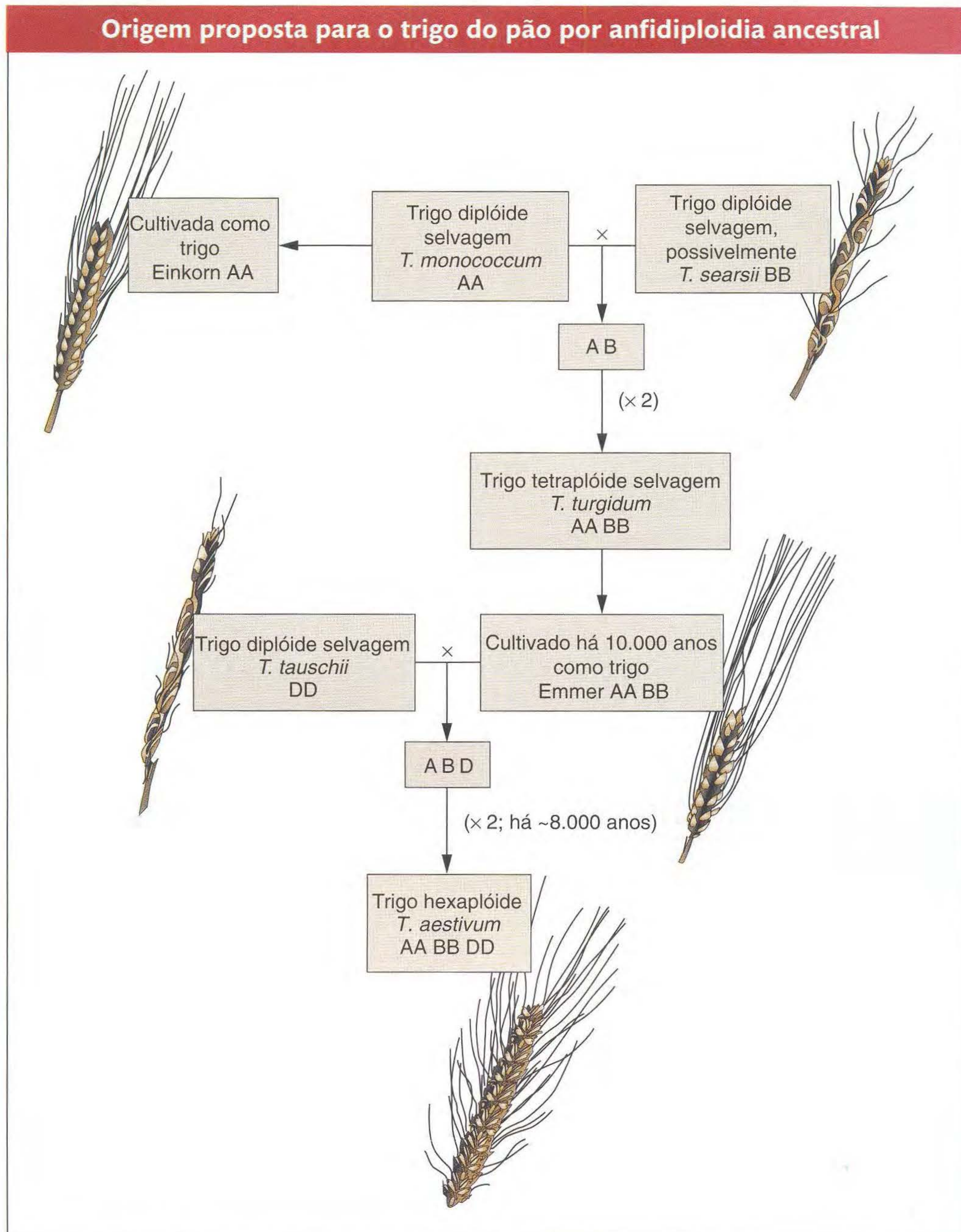


FIG. 16.9 Os anfidiplóides surgem em dois pontos na história evolutiva do moderno trigo hexaplóide. A, B e D são diferentes conjuntos cromossômicos.

usado comercialmente, um cultivo conhecido como *Triticale*. É um anfidiplóide entre trigo (*Triticum*, $6n = 42$) e centeio (*Secale*, $2n = 14$). Assim, para *Triticale*, $2n = 2 \times (21 + 7) = 56$. Essa nova planta combina a alta produtividade do trigo com a resistência da cevada.

Animais poliplóides Como já foi dito, a poliploidia é mais comum em plantas do que em animais, mas existem casos de ocorrência natural de animais poliplóides. Espécies poliplóides de platelmintos, sanguessugas e camarões de água salgada reproduzem-se por partenogênese. *Drosophila* triplóides e tetraplóides foram sintetizadas experimentalmente.

Entretanto, os exemplos não são limitados a essas chamadas formas inferiores. Anfíbios e répteis poliplóides de ocorrência natural são surpreendentemente comuns. Eles têm vários modos de reprodução: espécies poliplóides de sapos e girinos participam de reprodução sexual, enquanto salamandras e lagartos poliplóides são partenogenéticos. Os Salmonidae (a família de peixes que inclui o salmão e a truta) são um exemplo familiar das numerosas espécies animais que parecem ter se originado de poliploidia ancestral.

A esterilidade de triplóides foi comercialmente explorada tanto em animais como em plantas. Ostras triplóides foram desenvolvidas porque têm uma vantagem comercial em rela-

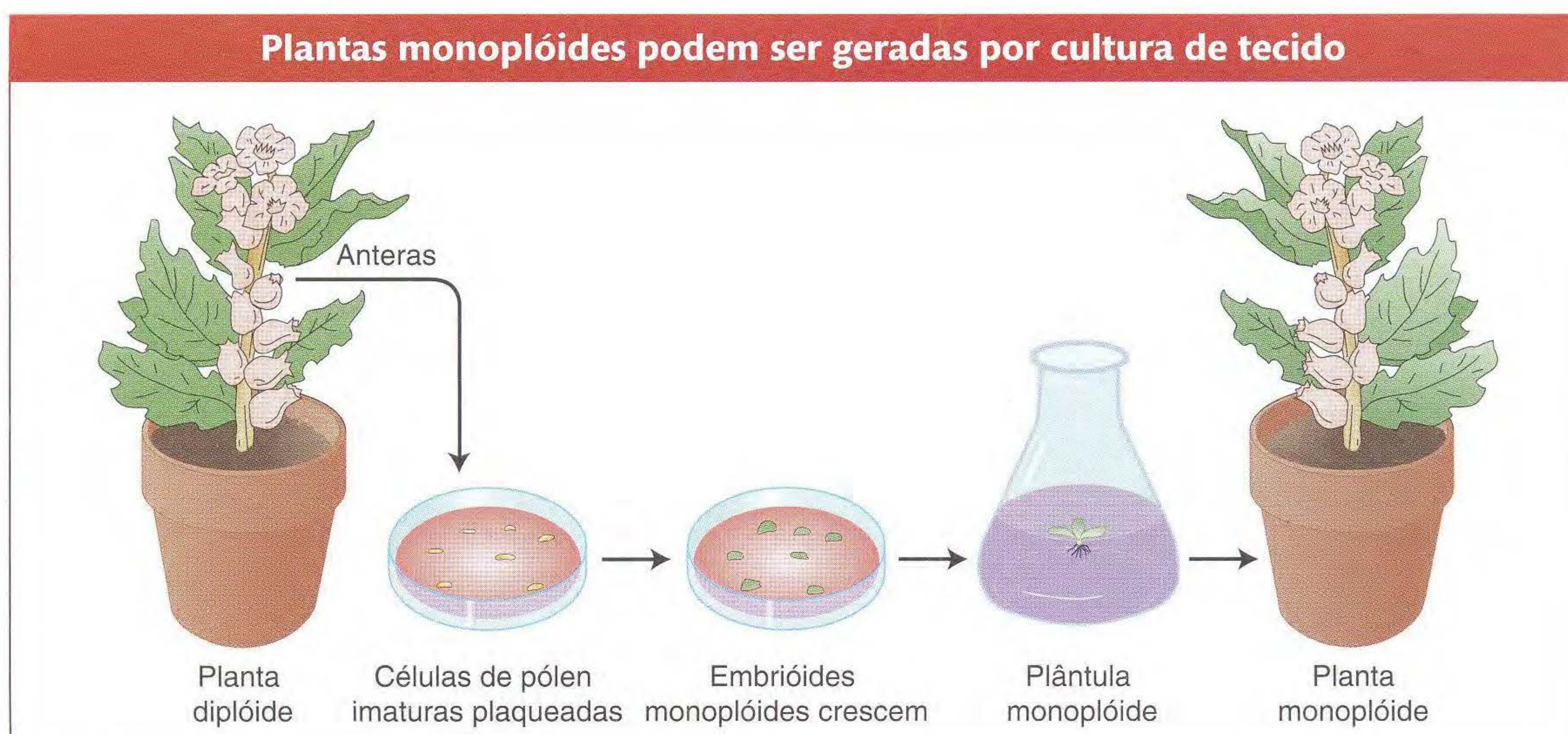


FIG. 16.10 Plantas monoplóides podem ser artificialmente derivadas de células destinadas a se tornarem grãos de pólen, expondo as células ao tratamento com frio em cultura de tecidos.

ção a seus parentes diplóides. Os diplóides passam por uma estação reprodutiva, quando não têm sabor bom, mas os triplóides estéreis não desovam e têm sabor o ano todo.

Aneuploidia

A aneuploidia é a segunda maior categoria de anomalias cromossômicas nas quais o número de cromossomos é anormal. Um aneuplóide é um organismo cujo número de cromossomos difere do tipo selvagem por parte de um conjunto de cromossomos. Em geral, o conjunto de cromossomos difere do tipo selvagem apenas por um cromossomo ou por um pequeno número de cromossomos. Um aneuplóide pode ter um número de cromossomos maior ou menor do que o tipo selvagem. A nomenclatura aneuplóide (veja Quadro 16.1) é baseada no número de cópias do cromossomo específico no estado aneuplóide. Para os autossomos em organismos diplóides, o aneuplóide $2n + 1$ é **trissômico**, $2n - 1$ é **monossômico**, e $2n - 2$ (o “-2” representa a perda de ambos os homólogos

de um cromossomo) é **nulissômico**. Nos haplóides, $n + 1$ é **dissômico**. A notação especial é usada para descrever os cromossomos sexuais aneuplóides porque deve lidar com os dois cromossomos diferentes. A notação meramente lista as cópias de cada cromossomo sexual, tais como XXY, XYY, XXX ou XO (o “O” indica a ausência de um cromossomo, e é incluído para mostrar que o único símbolo X não é um erro tipográfico).

Não-disjunção A causa da maioria das aneuploidias é a **não-disjunção** durante a meiose ou a mitose. *Disjunção* é outra palavra para a segregação normal de cromossomos homólogos ou cromátides para polos opostos na divisão meiótica ou mitótica. A *não-disjunção* é uma falha desse processo, na qual dois cromossomos, ou cromátides, incorretamente vão para um pólo e nenhum para o outro.

A não-disjunção *mitótica* pode ocorrer à medida que as células se dividem durante o desenvolvimento. Como resultado, trechos do corpo serão aneuplóides (*setores aneuplóides*). A não-disjunção *meiótica* é mais comumente encontrada. Nesse caso, os produtos da meiose são aneuplóides, levando a descendentes nos quais o organismo inteiro é aneuplóide. Na não-disjunção meiótica, os cromossomos podem falhar em se separar seja na primeira, seja na segunda divisão meiótica (Figura 16.12). Assim, são produzidos gametas $n - 1$ e $n + 1$. Se um gameta $n - 1$ é fertilizado por um gameta n , é produzido um zigoto monossômico ($2n - 1$). A fusão de um $n + 1$ e um gameta n produz um trissômico $2n + 1$.

Os organismos aneuplóides resultam principalmente de não-disjunção em uma meiose parental.

Uvas tetraplóides são maiores que as diplóides



FIG. 16.11 Uvas diplóides (esquerda) e tetraplóides (direita). [Copyright Leonard Lessin/Peter Arnold Inc.]

A não-disjunção ocorre espontaneamente. Como a maioria das mutações gênicas, é um exemplo de uma falha casual de um processo celular básico. Os processos moleculares exatos que falham não são conhecidos, mas, em sistemas experimen-

A não-disjunção meiótica gera produtos aneuplóides

W W W - ANIMAÇÃO COMPUTADORIZADA

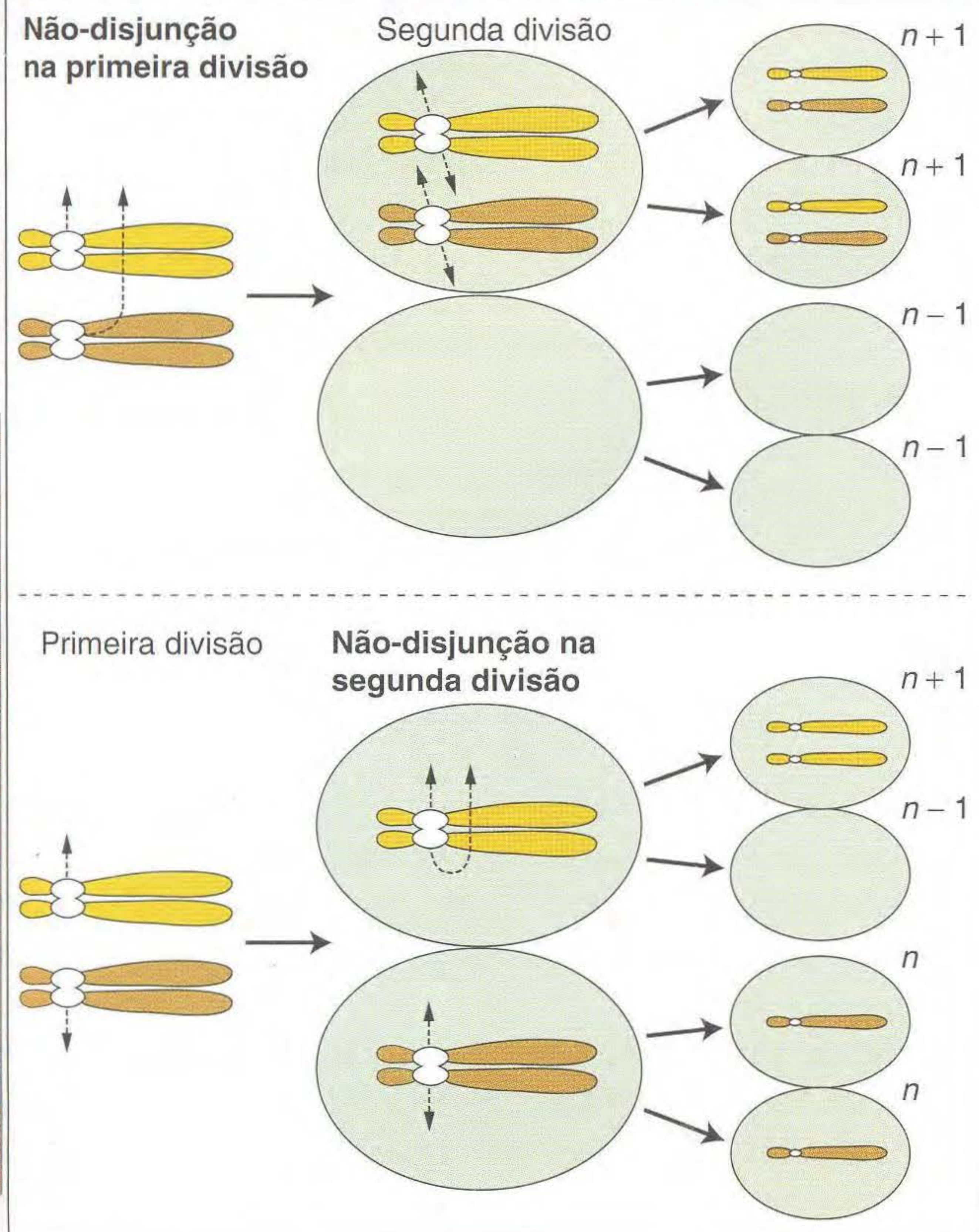


FIG. 16.12 Produtos aneuplóides da meiose (isto é, gametas) são produzidos por não-disjunção na primeira ou na segunda divisão meiótica. Note que todos os outros cromossomos estão presentes em número normal, incluindo as células nas quais não são mostrados os cromossomos.

tais, a frequência de não-disjunção pode ser aumentada pela interferência com a polimerização de microtúbulos, inibindo assim o movimento cromossômico normal. A não-disjunção parece ser mais comum na meiose I. Essa falha não é surpreendente porque a disjunção normal na anáfase I requer que as cromátides homólogas da tétrade permaneçam pareadas durante a prófase I e metáfase I, e ela requer os crossings. Em contraste, a disjunção apropriada na anáfase II ou na mitose requer que o centrômero divida-se apropriadamente, mas não requer o pareamento cromossômico ou crossing over.

Os crossings são um componente necessário do processo de disjunção normal. De algum modo, a formação de um quiasma ajuda a manter um bivalente junto e garante que as duas díades vão para polos opostos. Na maioria dos organismos, a quantidade de crossing over é suficiente para garantir que metade dos bivalentes tenham pelo menos um quiasma por meiose. Em *Drosophila*, muitos dos cromossomos não-disjuncionais vistos em gametas dissômicos ($n+1$) são não recombinantes, mostrando que eles surgem de meioses nas quais não há crossing over nesse cromossomo. Observações similares foram feitas em trissômicos humanos. Além disso, em vários organismos experimentais diferentes, as mutações

que interferem na recombinação têm o efeito de aumentar muito a frequência de não-disjunção na meiose I. Todas essas observações dão evidências do papel do crossing over em manter o pareamento cromossômico; na ausência dessas associações, os cromossomos são vulneráveis à não-disjunção na anáfase I.

Os crossings são necessários para manter os bivalentes pareados até a anáfase I. Se o crossing falha por algum motivo, ocorre uma não-disjunção na primeira divisão.

Monossômicos ($2n-1$) Os monossômicos não têm uma cópia de um cromossomo. Na maioria dos organismos diplóides, a ausência de uma cópia de um cromossomo de um par é deletéria. Em humanos, os monossômicos para qualquer dos autossomos morrem no útero. Muitos monossômicos para o cromossomo X também morrem no útero, mas alguns são viáveis. Um complemento de cromossomos humanos de 44 autossomos mais um único X produzem uma condição conhecida como **síndrome de Turner**, representada como XO. As pessoas afetadas têm um fenótipo característico: são mulheres estéreis, de baixa estatura, e, em geral, têm pescoço alado cuja extensão fica entre o pescoço e os ombros (Figura 16.13). Embora sua inteligência seja quase normal, algumas de suas funções cognitivas específicas são defeituosas. Cerca de 1 em cada 5.000 nascimentos femininos tem síndrome de Turner.

Os geneticistas têm usado plantas monossômicas viáveis para mapear alelos mutantes recessivos em um cromossomo específico. Por exemplo, podemos fazer um conjunto de

Características da síndrome de Turner (XO)

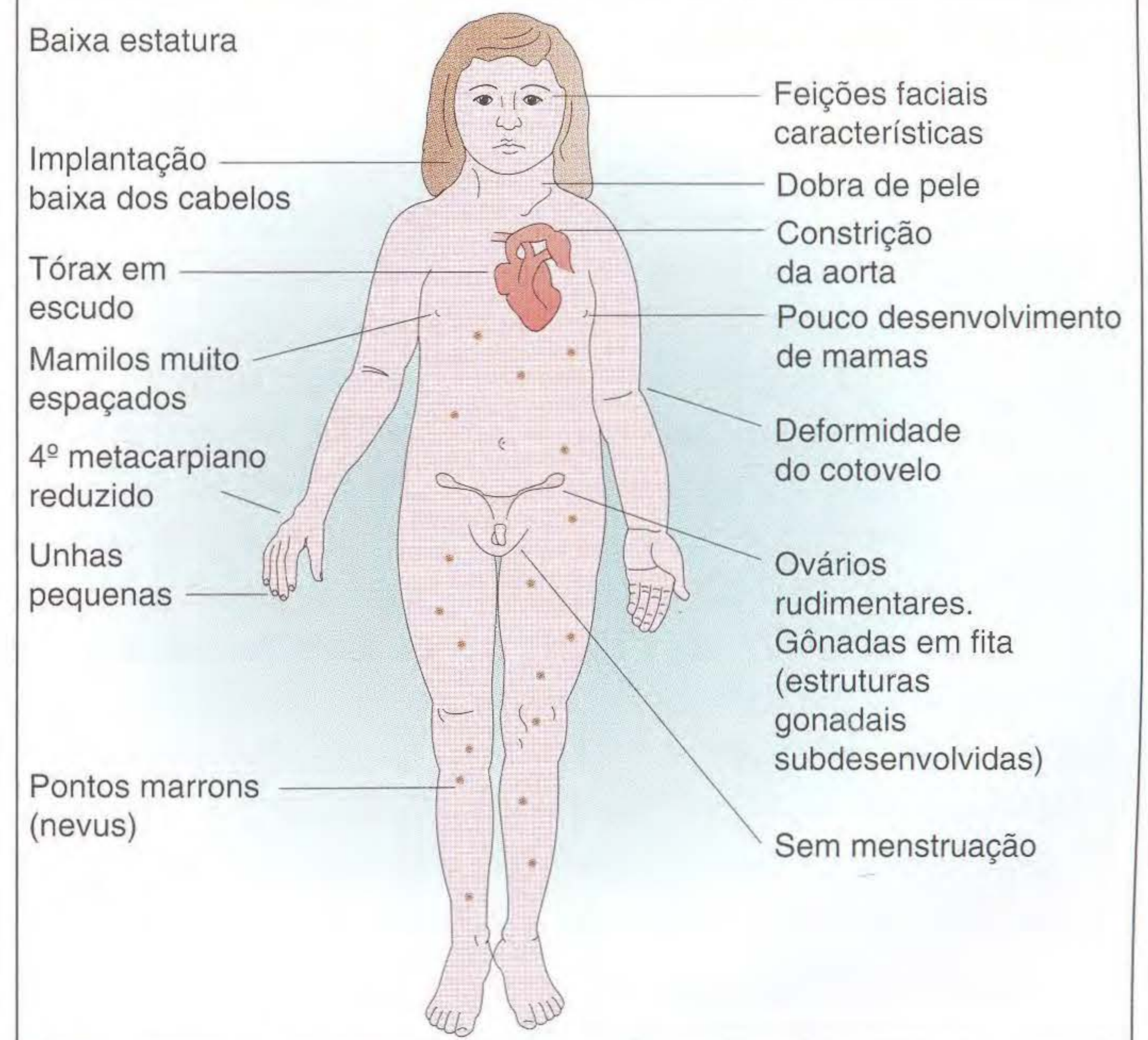


FIG. 16.13 A síndrome de Turner resulta da presença de um único cromossomo X (XO). [De F. Vogel and A.G. Motulsky, *Human Genetics*. Springer-Verlag, 1982.]

linhagens monossômicas, cada uma conhecida pela falta de um cromossomo diferente. Os homozigotos para novos alelos mutantes são cruzados com cada linhagem monossômica, e a prole de cada cruzamento é inspecionada quanto a um fenótipo recessivo. O surgimento de um fenótipo recessivo identifica o cromossomo que tem uma cópia faltando como aquela na qual o gene está normalmente situado. O teste funciona porque metade dos gametas de um monossômico fértil $2n - 1$ será $n - 1$, e, quando um gameta $n - 1$ é fertilizado por um gameta portador de uma nova mutação no cromossomo homólogo, o alelo mutante será o único alelo desse gene presente, e assim será expresso.

A título de ilustração, suponhamos que um gene A/a esteja no cromossomo 2. Os cruzamentos de a/a com monossômicos para os cromossomos 1 e 2 são pressupostos a fim de produzir diferentes resultados (o cromossomo 1 é abreviado como cro 1):

cro 1/cro 1 ; a/a Mutante	×	cro 1/0 ; A/A Cromossomo 1 monossômico genótipo A
↓		
prole		toda A/a
cro 1/cro 1 ; a/a Mutante	×	cro 1/cro 1 ; $A/0$ Cromossomo 2 monossômico genótipo A
↓		
prole		$\frac{1}{2} A/a$ $\frac{1}{2} a/0$

Trissômicos ($2n + 1$) Os trissômicos contêm uma cópia extra de um cromossomo. Em organismos diplóides, geralmente o desequilíbrio cromossômico da condição trissômica pode resultar em anomalia ou morte. Entretanto, existem muitos exemplos de trissômicos viáveis. Além disso, os trissômicos podem ser férteis. Quando as células de alguns organismos trissômicos são observadas ao microscópio na época do pareamento meiótico, os cromossomos trissômicos são vistos formando um grupo associado de três (trivalente), enquanto os outros cromossomos formam bivalentes regulares.

Que proporções genéticas podem ser esperadas para os genes no cromossomo trissômico? Consideremos um gene A que está próximo do centrômero nesse cromossomo, e vamos supor que o genótipo é $A/a/a$. Além disso, vamos postular que, na anáfase I, os dois centrômeros pareados no trivalente passam para pólos opostos e que o outro centrômero passe aleatoriamente para qualquer um dos pólos. Então podemos prever as três segregações igualmente freqüentes mostradas na Figura 16.14. Essas segregações resultam em uma proporção gamética como as mostradas nos seis compartimentos da Figura 16.14; isto é,

$$\begin{array}{l} \frac{1}{6} A \\ \frac{2}{6} a \\ \frac{2}{6} A/a \\ \frac{1}{6} a/a \end{array}$$

Produtos meióticos de um trissômico		
	Um pólo	Outro pólo
	1 A 2 a	3 a
	1 A 3 a	2 a
	2 a 3 a	1 A

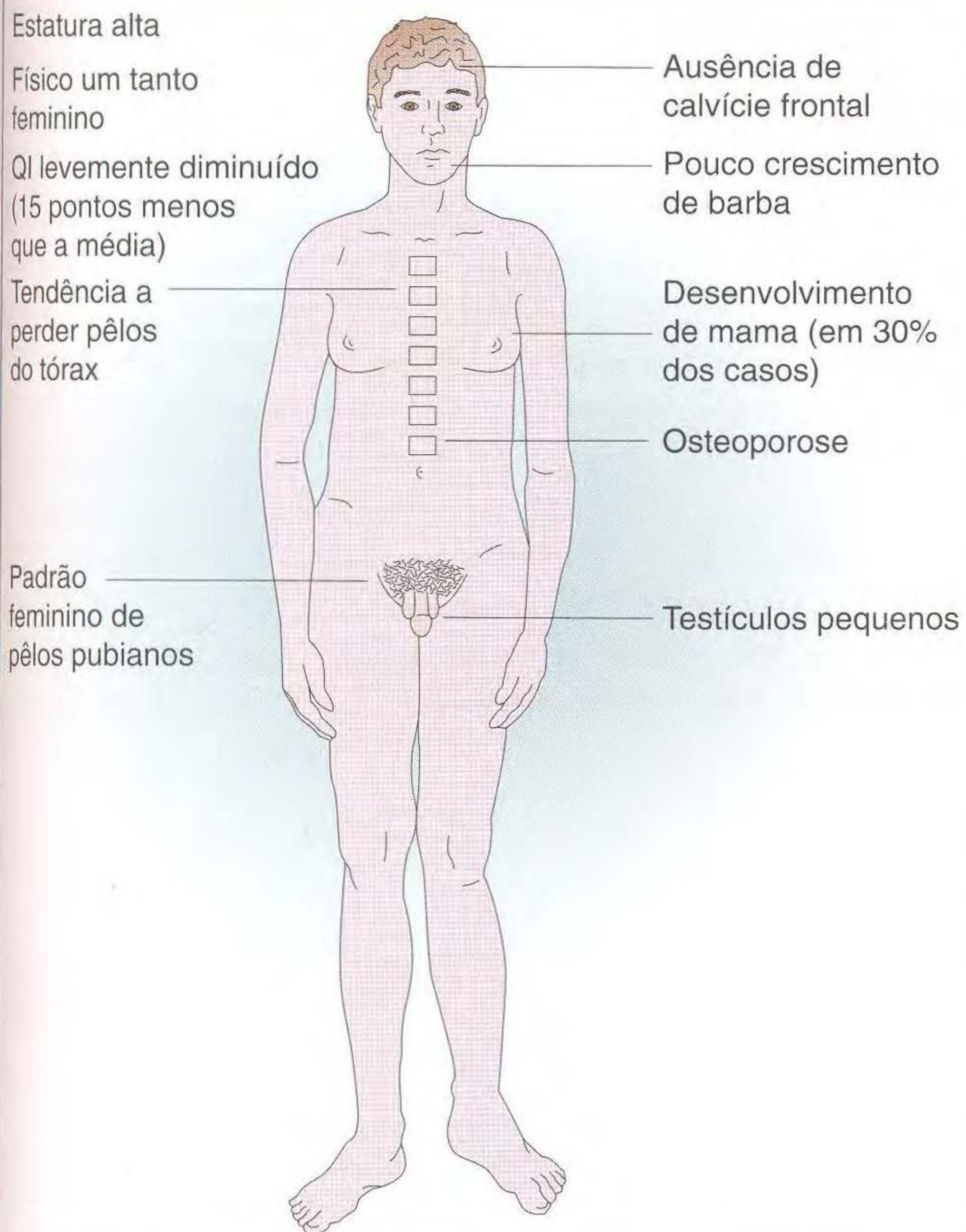
FIG. 16.14 Três segregações igualmente prováveis podem ocorrer na meiose de um trissômico $A/a/a$, contendo os genótipos mostrados.

Se uma série de linhagens estiver disponível, cada uma levando um cromossomo trissômico diferente, então uma mutação gênica pode ser situada em um cromossomo, determinando qual das linhagens dá uma proporção trissômica do tipo precedente.

Existem vários exemplos de trissômicos humanos viáveis. Vários tipos de trissômicos de cromossomos sexuais podem viver até a idade adulta. Cada um desses tipos é encontrado com uma freqüência de cerca de 1 em 1.000 nascimentos do sexo relevante. (Considerando as trissomias de cromossomos sexuais humanos, lembre que o sexo de mamíferos é determinado pela presença ou ausência do cromossomo Y.) A combinação XXY resulta na **síndrome de Klinefelter**. As pessoas com essa síndrome são homens que têm uma constituição magra, um QI levemente prejudicado e são estéreis (Figura 16.15). Outra combinação anormal, XYY, tem uma história controversa. Foram feitas tentativas de ligar a condição XYY a uma predisposição para a violência. Entretanto, hoje está claro que uma condição XYY de modo algum garante tal comportamento. A maioria dos homens XYY são férteis. A meiose apresenta pareamento normal do X com um dos Y; o outro Y não pareia e não é transmitido aos gametas. Portanto, os gametas contêm ou X ou Y, nunca YY ou XY. Os trissômicos triplo X (XXX) são fenotipicamente normais e mulheres férteis. A meiose mostra o pareamento apenas de dois cromossomos X; o terceiro não pareia. Assim, os ovócitos têm apenas um X e, como nos homens XYY, a condição não é transmitida para a prole.

Das trissomias humanas, o tipo mais comum é a **síndrome de Down** (Figura 16.16), discutida resumidamente no começo do capítulo. A freqüência da síndrome de Down é de cerca de 0,15% de todos os nativos. A maioria das pessoas afetadas tem uma cópia extra do cromossomo 21 causada por não-disjunção do cromossomo 21 em um genitor que é cromossomicamente normal. Nesse tipo *esporádico* de síndrome de Down, não há história familiar de aneuploidia. Alguns tipos raros de síndrome de Down surgem de translocações (um tipo de rearranjo cromossômico discutido mais adiante no capítulo); nesses casos, como veremos, a síndrome de

Características da síndrome de Klinefelter (XXY)



Down ocorre novamente no heredograma porque a translocação pode ser transmitida de genitor ao filho.

Os fenótipos combinados que constituem a síndrome de Down incluem retardo mental (com um QI na faixa de 20 a 50); uma face larga e achatada, olhos com pregas epicânticas, baixa estatura, mãos curtas com um sulco no meio; e uma língua grande e com rugas. As mulheres podem ser férteis e ter uma prole normal ou trissômica, mas os homens são estéreis com poucas exceções. A expectativa média de vida é de cerca de 17 anos, e apenas 8% das pessoas com síndrome de Down sobrevivem após os 40 anos.

A incidência da síndrome de Down está relacionada à idade materna; mães com mais idade correm um risco elevado de síndrome de Down (Figura 16.17). Por esse motivo, a análise cromossômica (por amniocentese ou por amostra de vilosidades coriônicas) hoje é recomendada para mães com mais idade. Um efeito menos pronunciado de idade paterna também foi demonstrado.

Muito embora o efeito da idade materna seja conhecido há muitos anos, sua causa ainda é desconhecida. Entretanto, existem algumas correlações biológicas interessantes. Com a idade, possivelmente, torna-se menos provável que os bivalentes cromossômicos permaneçam juntos durante a prófase I da meiose. A parada meiótica dos ovócitos (meiócitos femininos) na prófase I avançada é um fenômeno comum em muitos animais. Nas mulheres, todos os meiócitos estão parados no diplóteno antes do nascimento. A meiose retorna a cada período menstrual, o que significa que os cromossomos no bivalente devem permanecer apropriadamente associados por pelo menos cinco ou mais décadas. Se especularmos que essas associações têm um aumento de probabilidade de se romper por acidente com o passar do tempo, podemos imaginar um mecanismo que contribua para o aumento da

FIG. 16.15 A síndrome de Klinefelter resulta da presença de dois cromossomos X e um cromossomo Y. [De F. Vogel and A.G. Motulsky, *Human Genetics*. Springer-Verlag, 1982.]

Características da síndrome de Down (trissomia do 21)

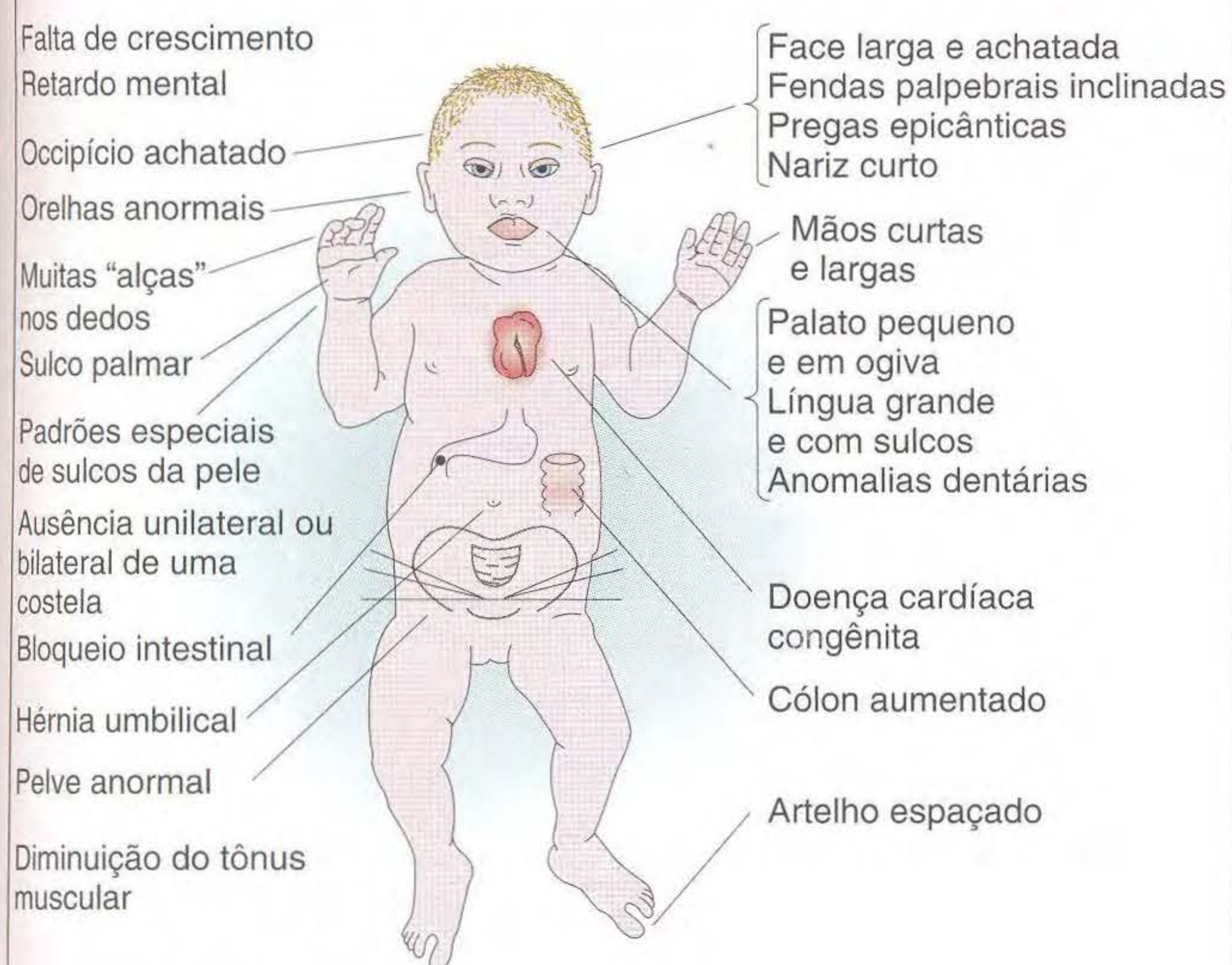


FIG. 16.16 A síndrome de Down resulta da presença de uma cópia extra do cromossomo 21. [De F. Vogel and A.G. Motulsky, *Human Genetics*. Springer-Verlag, 1982.]

Efeito da idade materna na síndrome de Down

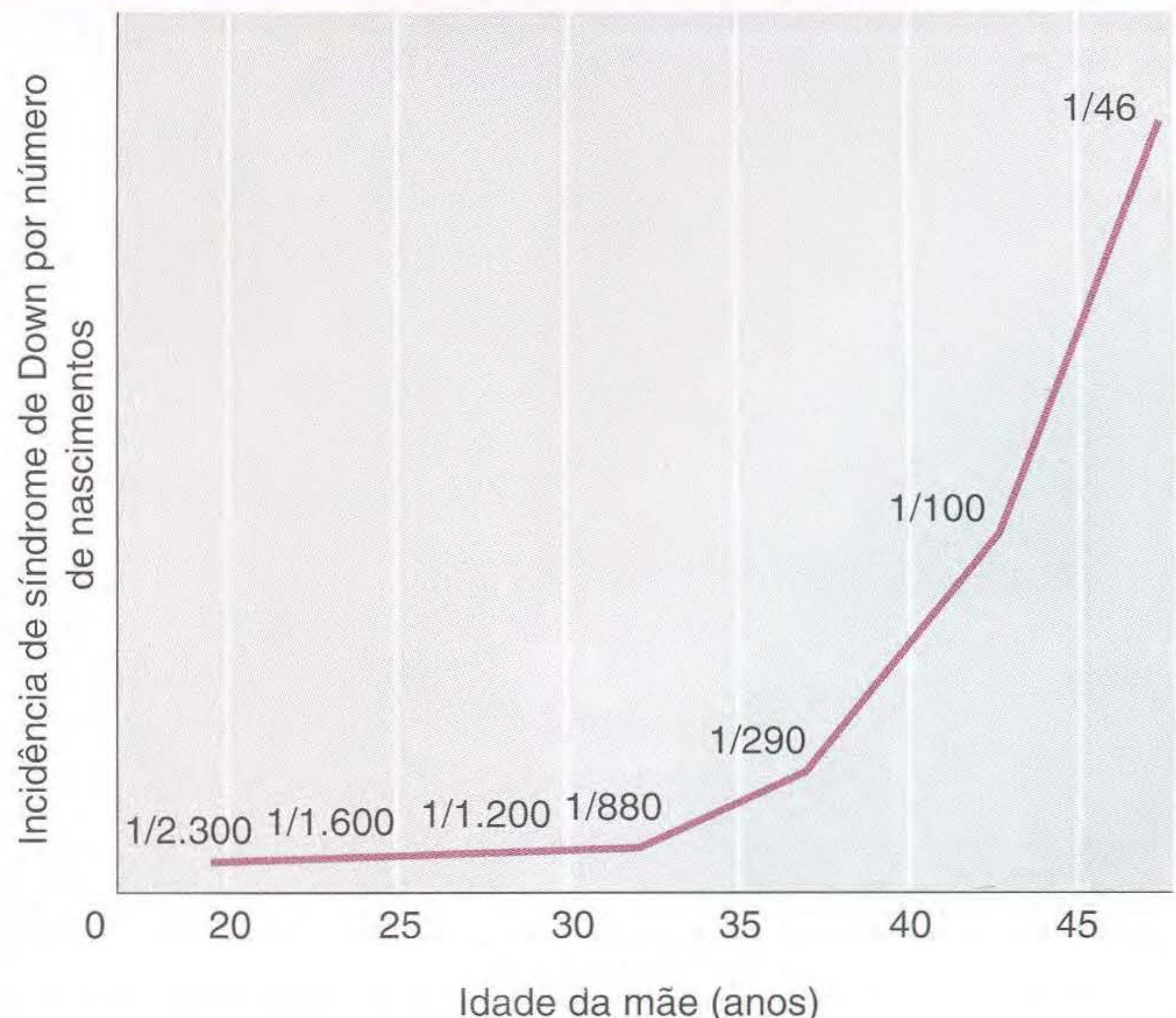


FIG. 16.17 Mulheres com mais idade têm uma proporção maior de crianças com síndrome de Down do que mulheres mais jovens. [De L.S. Penrose and G.F. Smith, *Down's Anomaly*. Little, Brown and Company, 1966.]

não-disjunção materna com a idade. Consistente com essa especulação, a maioria das não-disjunções relacionadas ao efeito da idade materna é devida à não-disjunção na anáfase I, e não na anáfase II.

As únicas outras trissomias autossômicas humanas a sobreviverem até o nascimento são a trissomia do 13 (síndrome de Patau) e a trissomia do 18 (síndrome de Edwards). Ambas têm graves anomalias físicas e mentais. A síndrome fenotípica da trissomia do 13 inclui fenda labial; uma cabeça pequena e malformada; pés “em cadeira de balanço”; e uma expectativa média de vida de 130 dias. A trissomia do 18 inclui orelhas “tipo fauno”, mandíbula pequena, pelve estreita e pés “de cadeira de balanço”. Quase todas as crianças com trissomia do 18 morrem dentro das primeiras semanas após o nascimento. Todas com as outras trissomias morrem no útero.

O conceito de balanço gênico

Considerando as euploidias aberrantes, notamos que uma incidência no número de todos os conjuntos cromossômicos está correlacionado ao tamanho do organismo, mas que a forma geral e as proporções do organismo permanecem praticamente as mesmas. Em contraste, a aneuploidia autossômica altera tipicamente a forma do organismo e as proporções de modos característicos.

As plantas tendem a ser um pouco mais tolerantes à aneuploidia do que os animais. Os estudos em estramônio (*Datura stramonium*) dão um exemplo clássico dos efeitos da aneuploidia e poliploidia. Em estramônio, o número haplóide de cromossomos é 12. Como esperado, a poliploidia de estramônio é proporcionada como a diploidia normal, apenas maior. Em contraste, cada uma das 12 possíveis trissomias é desproporcionada, mas de modos diferentes umas das outras,

como exemplificado pelas mudanças na forma da cápsula da semente (Figura 16.18). As 12 trissomias diferentes levam a 12 mudanças de forma diferentes e características na cápsula. De fato, essas características e outras trissomias individuais são tão confiáveis que a síndrome fenotípica pode ser usada para identificar plantas levando determinada trissomia. Similarmente, os 12 monossômicos são diferentes uns dos outros e de cada um dos trissômicos. Em geral, um monossômico para determinado cromossomo é mais gravemente anormal do que o trissômico correspondente.

Vemos tendências similares em animais aneuplóides. Na mosca das frutas *Drosophila*, únicos aneuplóides autossômicos que sobrevivem até a vida adulta são trissômicos e monossômicos para o cromossomo 4, que é o menor cromossomo de *Drosophila*, representando apenas cerca de 1 a 2% do genoma. Os trissômicos para o cromossomo 4 são apenas brandamente afetados e são muito menos anormais do que os monossômicos para o cromossomo 4. Em humanos, nenhum autossomo monossômico sobrevive até o nascimento, mas, como já foi dito, três tipos de trissomias autossômicas podem fazer isso. Como é verdade para o estramônio aneuplóide, cada um dos três trissômicos mostra síndromes fenotípicas únicas devido aos efeitos especiais de dosagens aumentadas de cada um desses cromossomos.

Por que os aneuplóides são muito mais anormais que os poliplóides? Por que a aneuploidia para cada cromossomo tem seus próprios efeitos fenotípicos característicos? E por que os monossômicos são mais gravemente afetados que os trissômicos correspondentes? As respostas parecem ser uma questão de **balanço gênico**. Em um euplóide, a proporção de genes de qualquer cromossomo em relação aos genes no outro cromossomo é sempre de 1:1, independentemente de estarmos considerando monoplóide, diplóide, triplóide ou

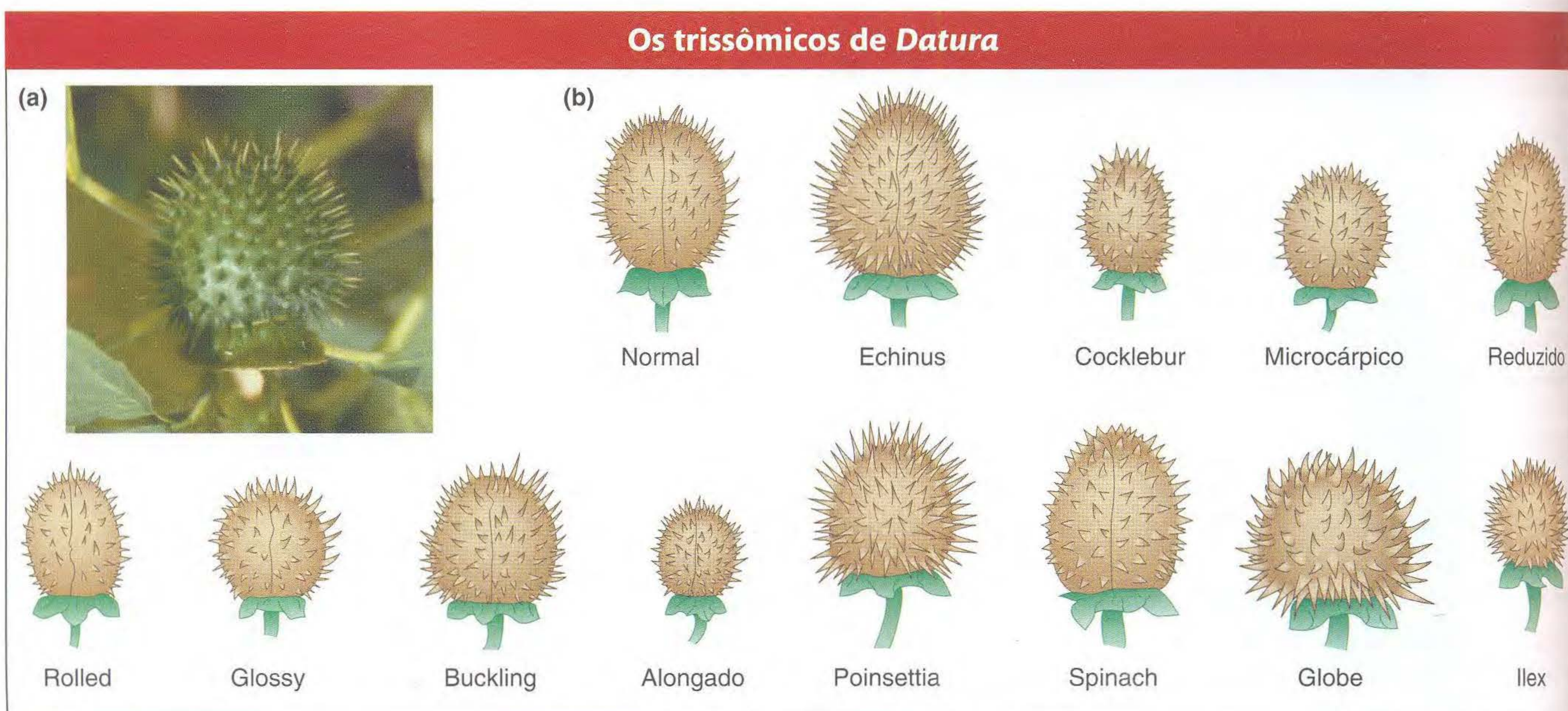


FIG. 16.18 Cada um dos 12 possíveis trissômicos de *Datura* é desproporcional de um modo diferente. (a) Fruto de *Datura*. (b) Cada desenho é do fruto de um trissômico diferente, cada um dos quais foi nomeado. [(a) Dr. G.W.M. Barendse/Nijmegen University Botanical Garden; (b) Adaptado de F.W. Sinnott, L.C. Dunn, and T. Dobzhansky, *Principles of Genetics*, 5th ed. McGraw-Hill Book Company, 1958.]

tetraplóide. Por exemplo, em um tetraplóide, para o gene *A* no cromossomo 1 e gene *B* no cromossomo 2, a proporção é 4 *A*:4 *B*, ou 1:1. Em contraste, em um aneuplóide, a proporção de genes no cromossomo aneuplóide para genes no outro cromossomo difere do tipo selvagem em 50%: 50% de monossômicos; 150% para trissômicos. Usando o mesmo exemplo que antes, em um trissômico para o cromossomo 2, encontramos que a proporção de genes *A* e *B* é 2 *A*:3 *B*. Assim, podemos ver que os genes aneuplóides estão desbalanceados. Como não estarem balanceados pode nos ajudar a responder as perguntas levantadas?

Em geral, a quantidade de transcritos produzidos por um gene é diretamente proporcional ao número de cópias desse gene em uma célula. Isto é, para determinado gene, a taxa de transcrição está diretamente relacionada ao número de moldes de DNA disponíveis. Assim, quanto mais cópias do gene, mais transcritos são produzidos e mais é feito o produto proteico correspondente. Essa relação entre o número de cópias de um gene e a quantidade do produto gênico é chamada de **efeito de dosagem gênica**.

Podemos deduzir que a fisiologia normal em uma célula depende da proporção apropriada de produtos gênicos na célula euplóide. Essa proporção é o balanço gênico normal. Se a dosagem relativa de determinado gene muda, por exemplo, devido à remoção de uma das duas cópias de um cromossomo (ou mesmo de um segmento), podem surgir desequilíbrios fisiológicos em vias celulares.

Em alguns casos, os desequilíbrios de aneuploidia resultam de efeitos de alguns genes “importantes” cuja dosagem mudou, e não de mudanças na dosagem de todos os genes em um cromossomo. Tais genes podem ser vistos como *haplo-anormais* (resultando em um fenótipo anormal se presente apenas uma vez) ou *triplo-anormais* (resultando em um fenótipo anormal se presente em três cópias) ou ambos. Eles contribuem significativamente para as síndromes fenotípicas aneuplóides. Por exemplo, o estudo de pessoas trissômicas apenas para parte do cromossomo 21 possibilitou localizar os genes que contribuem para a síndrome de Down em várias regiões do cromossomo 21. Os resultados indicam que alguns aspectos do fenótipo podem ser devidos a tripla anomalia de genes únicos importantes nessas regiões cromossômicas. Além desses importantes efeitos gênicos, outros aspectos das síndromes de aneuploidia provavelmente resultam de efeitos cumulativos de aneuploidia para vários genes cujos produtos estão todos desbalanceados. Sem dúvida, todo o fenótipo aneuplóide resulta de uma combinação dos efeitos de desbalanço de alguns genes importantes, junto com um desequilíbrio cumulativo de muitos genes secundários.

Entretanto, o conceito de balanço gênico não nos diz por que ter muito pouco produto gênico (monossomia) é muito pior para um organismo do que ter excesso de produtos gênicos (trissomia). De um modo paralelo, podemos perguntar por que existem muito mais genes haplo-anormais do que triplo-anormais. Uma chave para explicar a extrema anomalia dos monossômicos é que quaisquer alelos recessivos deletérios presentes em um autossomo monossômico será automaticamente expresso.

Como aplicamos a idéia de balanço gênico a casos de aneuploidia de cromossomos sexuais? O balanço gênico vale também para cromossomos sexuais, mas também temos que levar em conta as propriedades especiais dos cromossomos sexuais. Em organismos com determinação sexual XY, o cromossomo Y parece ser um cromossomo X degenerado no qual existem poucos genes funcionais além dos envolvidos na própria determinação do sexo, na produção de espermatozoides ou em ambos. O cromossomo X, por outro lado, contém muitos genes envolvidos em processos celulares básicos (“genes de manutenção”) residentes no cromossomo que, ocasionalmente, evoluiu no cromossomo X. Os mecanismos de determinação XY do sexo provavelmente evoluíram, independentemente, de 10 a 20 vezes em grupos taxonômicos diferentes. Por exemplo, parece haver um mecanismo de determinação do sexo para todos os mamíferos, mas é completamente diferente do mecanismo que controla a determinação XY do sexo na mosca-das-frutas.

De certo modo, os cromossomos X são naturalmente aneuplóides. Nas espécies com um sistema de determinação do sexo XY, as fêmeas têm dois cromossomos X, enquanto os machos têm apenas um. Entretanto, os genes de manutenção do cromossomo X são expressos de modo aproximadamente igual por célula nas fêmeas e nos machos. Em outras palavras, há uma **compensação de dose**. Como é estabelecida essa compensação? A resposta depende do organismo. Nas moscas-das-frutas, o cromossomo X dos machos parece ter sido hiperativado, deixando que ele seja transcrito no dobro da taxa de qualquer cromossomo X na fêmea. Como resultado, o macho XY de *Drosophila* tem uma dosagem do gene X equivalente à de uma fêmea XX. Nos mamíferos, em contraste, a regra é que não importa quantos cromossomos X estejam presentes, existe apenas um cromossomo X transcricionalmente ativo em cada célula somática. Essa regra dá à fêmea XX de mamífero uma dosagem gênica de X equivalente à de um macho XY. A compensação de dose em mamíferos é obtida pela inativação do cromossomo X. Uma fêmea com dois cromossomos X, por exemplo, é um mosaico de dois tipos de células nas quais um ou outro cromossomo X é ativo. Examinamos esse fenômeno no Capítulo 11. Assim, os indivíduos XY e XX produzem as mesmas quantidades de genes de manutenção do cromossomo X. A inativação do cromossomo X também explica por que os humanos triplo X são fenotipicamente normais: apenas um dos três cromossomos X é transcricionalmente ativo em determinada célula. Similarmente, um homem XXY é apenas moderadamente afetado porque apenas um de seus dois cromossomos X é ativo em cada célula.

Por que os indivíduos XXY são anormais, considerando que os indivíduos triplo X são fenotipicamente normais? Ocorre que alguns genes distribuídos por um X “inativo” ainda são transcricionalmente ativos. Nos homens XXY, esses genes são transcritos no dobro do nível que são nos homens XY. Nas mulheres XXX, por outro lado, os poucos genes transcritos são ativos apenas 1,5 vez o nível que são nas mulheres XX. Esse nível menor de “aneuploidia funcional” em XXX do que em XXY mais o fato de que os genes X ativos parecem levar

à feminização podem explicar o fenótipo feminizado dos machos XXY. A gravidade da síndrome de Turner (XO) pode ser devida aos efeitos deletérios de monossomia e à menor atividade dos genes transcritos de mulheres X (comparados com XX). Como é geralmente observado em aneuplóides, a monossomia do cromossomo X produz um fenótipo mais anormal do que ter uma cópia extra do mesmo cromossomo (mulheres triplo X ou homens XXY).

A dosagem gênica também é importante nos fenótipos poliplóides. Os zigotos poliplóides humanos surgem por vários tipos de erros na divisão celular. A maioria morre no útero. Ocasionalmente, nascem crianças triplóides, mas não sobrevivem. Esse fato parece violar o princípio de que poliplóides são mais normais que os aneuplóides. A explicação para essa contradição parece estar na compensação de dose. Parte da regra para o balanço gênico, nos organismos que têm um complemento cromossômico de um único X ativo, parece ser que deve haver um único X ativo para cada duas cópias do complemento cromossômico autossômico. Assim, algumas células em mamíferos triplóides são vistas tendo um X ativo, enquanto outras, surpreendentemente, têm dois. Nenhuma dessas situações está em balanço com genes autossômicos.

A aneuploidia é quase sempre deletéria devido ao desbalanceamento gênico: a proporção de genes é diferente da observada em euplóides, e essa diferença interfere na função normal do genoma.

► 16.2 Mudanças na Estrutura de Cromossomos

As mudanças na estrutura de cromossomos, chamadas de **rearranjos**, englobam várias classes importantes de eventos. Um segmento cromossômico pode ser perdido, constituindo uma **deleção**, ou duplicado para formar uma **duplicação**. A orientação de um segmento dentro do cromossomo pode ser revertida, constituindo uma **inversão**. Um segmento pode ser movido para um cromossomo diferente, constituindo uma **translocação**. A quebra do DNA é uma causa importante de cada um desses eventos. Ambos os filamentos do DNA podem ser quebrados em dois locais diferentes, seguidos de reunião das pontas quebradas para produzir um novo arranjo cromossômico (Figura 16.19, esquerda). Os rearranjos cromossômicos por quebras podem ser induzidos artificialmente usando radiação ionizante. Esse tipo de radiação, particularmente raios X e raios gama, é de alta energia e causa várias quebras bifilamentares no DNA.

Para compreendermos como são produzidos os rearranjos cromossômicos por quebras, devemos ter em mente vários pontos:

1. Cada cromossomo é uma única molécula bifilamentar de DNA.
2. O primeiro evento na produção de um rearranjo cromossômico é a geração de duas ou mais quebras bifilamentares

nos cromossomos de uma célula (veja Figura 16.19, fileira de cima à esquerda).

3. As quebras bifilamentares são potencialmente letais, a menos que sejam reparadas.
4. Os sistemas de reparo na célula corrigem as quebras bifilamentares juntando pontas quebradas (veja o Capítulo 15 para uma discussão detalhada de reparo do DNA).
5. Se as duas pontas da mesma quebra forem reunidas, a ordem original do DNA é restaurada. Se as pontas de duas quebras diferentes forem unidas, entretanto, um resultado é um tipo ou outro de rearranjo cromossômico.
6. Os únicos rearranjos cromossômicos que sobrevivem à meiose são os que produzem moléculas de DNA que têm um centrômero e dois telômeros. Se um rearranjo produz um cromossomo que não tem centrômero, tal cromossomo **acêntrico** não será levado para nenhum dos pólos na anáfase da mitose ou meiose, e não será incorporado a nenhum núcleo da prole. Assim, os cromossomos acêntricos não são herdados. Se um rearranjo produz um cromossomo com dois centrômeros (um **dicêntrico**), ele será arrastado simultaneamente para pólos opostos na anáfase, formando uma **ponte anafásica**. Cromossomos de ponte anafásica tipicamente não serão incorporados a nenhuma célula-filha. Se uma quebra cromossômica produz um cromossomo sem um telômero, esse cromossomo não pode replicar-se apropriadamente. Lembre, do Capítulo 7, que os telômeros são necessários para iniciar a replicação apropriada do DNA nas pontas (veja Figura 7.25).
7. Se um rearranjo duplica ou deleta um segmento de um cromossomo, o balanço gênico pode ser afetado. Quanto maior o segmento que é perdido ou duplicado, mais provável que desequilíbrio gênico cause anomalias fenotípicas.

Outra causa importante de rearranjo é o crossing over entre segmentos de DNA repetitivos (duplicados). Esse tipo de crossing over é chamado de **recombinação homóloga não-alélica (NAHR)**. Nos organismos com seqüências repetidas de DNA dentro de um cromossomo ou em cromossomos diferentes, há uma ambigüidade sobre qual das repetições fará par uma com a outra na meiose. Se as seqüências que pareiam não estão nas mesmas posições relativas dos homólogos, o crossing over pode produzir cromossomos aberrantes. Podem ser produzidas deleções, duplicações, inversões e translocações por tal crossing over (veja Figura 16.19, lado direito).

Existem dois tipos gerais de rearranjos: desbalanceados e balanceados. Os **rearranjos desbalanceados** mudam a dosagem gênica de um segmento cromossômico. Como na aneuploidia para cromossomos inteiros, a perda de uma cópia de um segmento ou a adição de uma cópia extra podem perturbar o balanço gênico normal. As duas classes simples de rearranjos desbalanceados são deleções e duplicações. Uma **deleção** é a perda de um segmento dentro de um braço cromossômico e a justaposição dos dois segmentos em ambos os lados do segmento deletado, como neste exemplo, que mostra a perda do segmento C-D:

Origens dos rearranjos cromossômicos

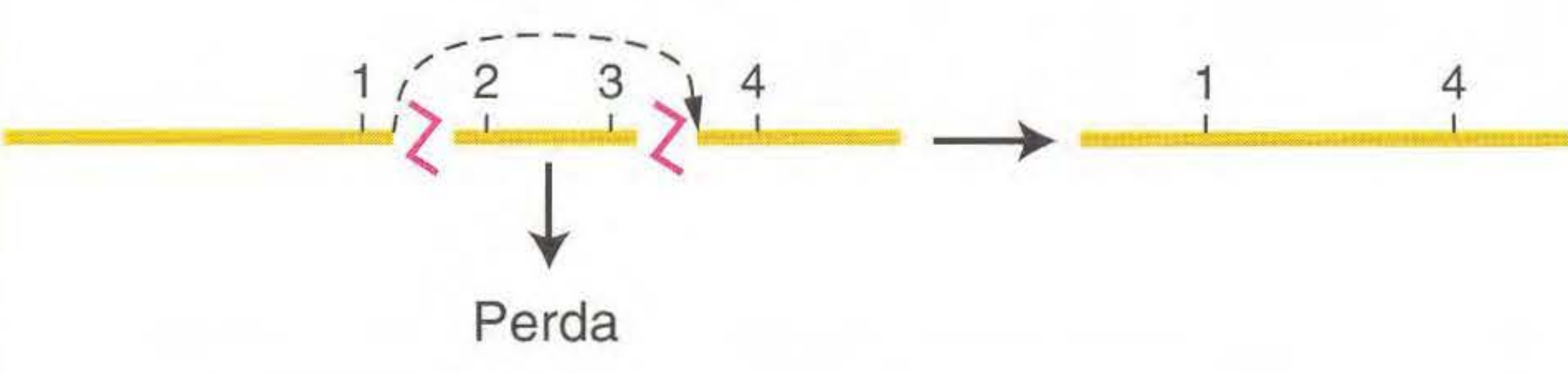
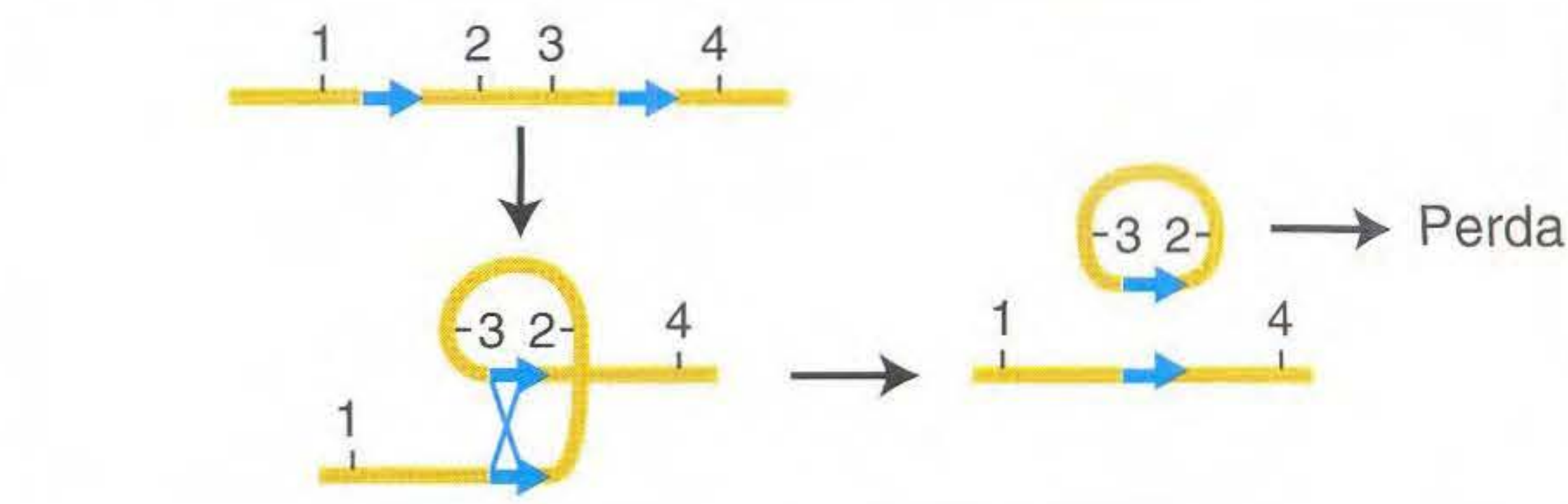
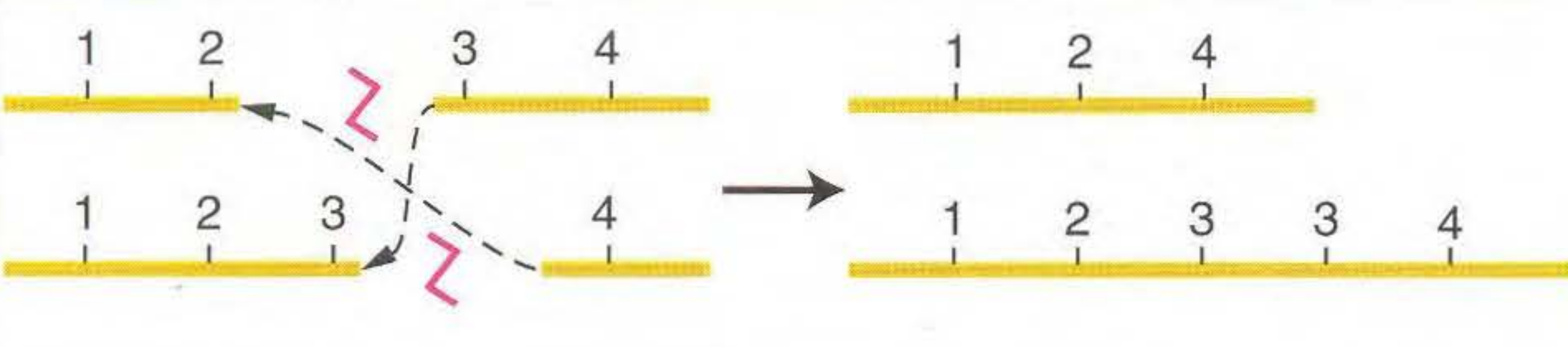
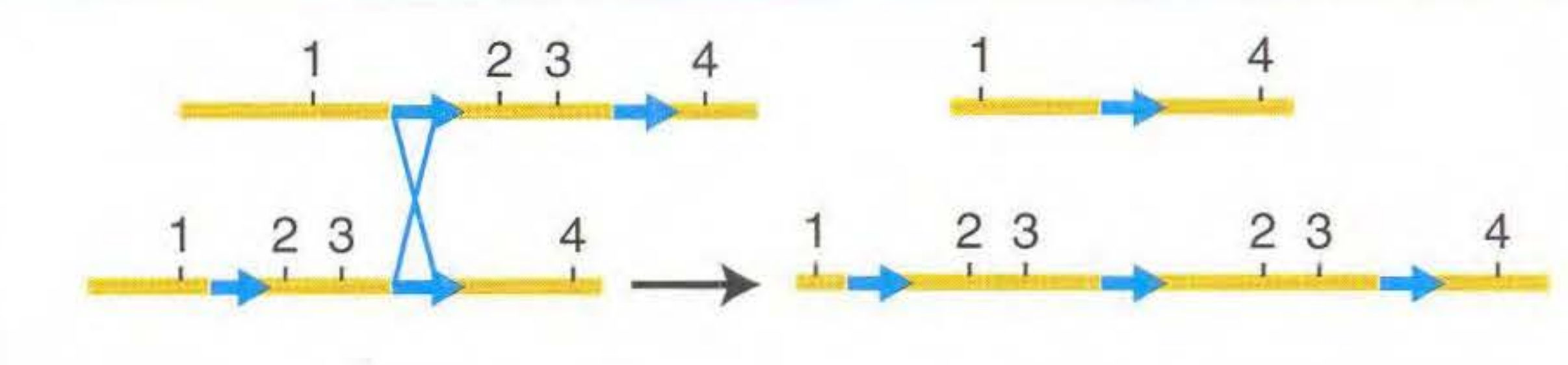
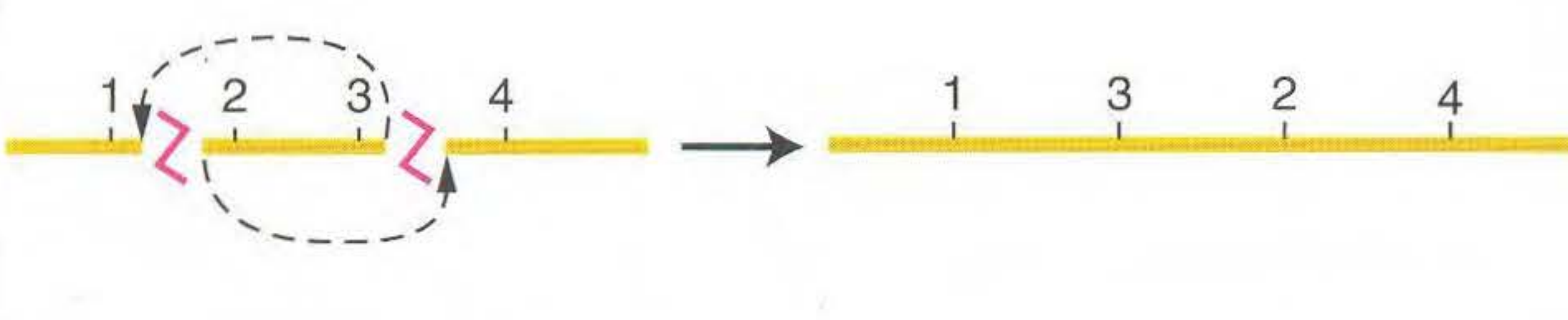
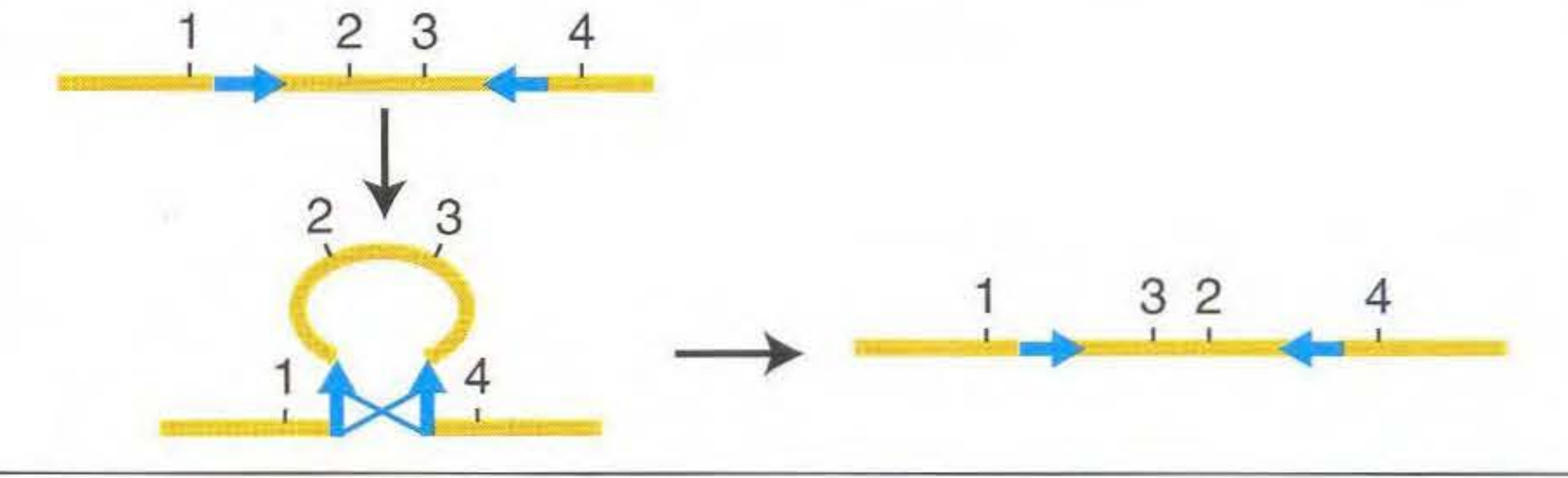
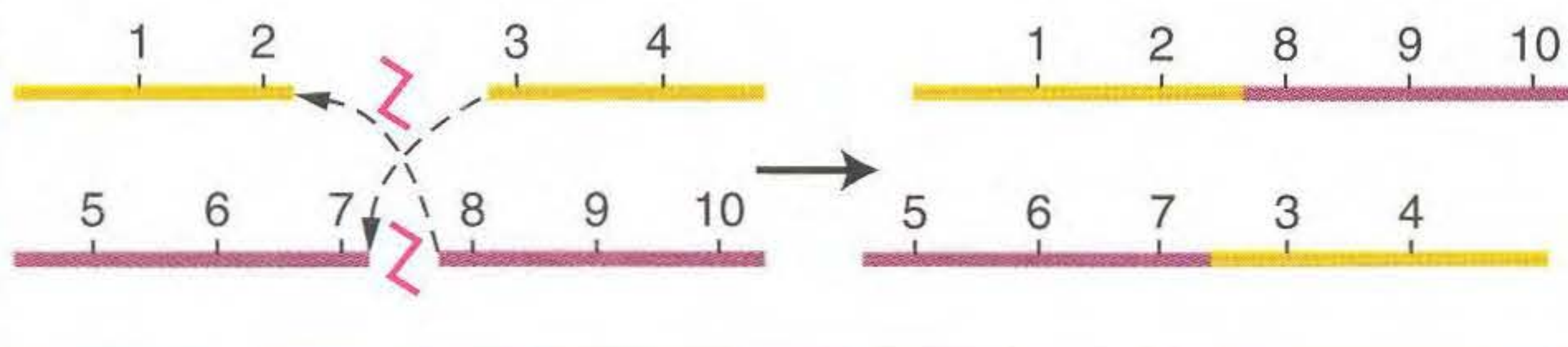
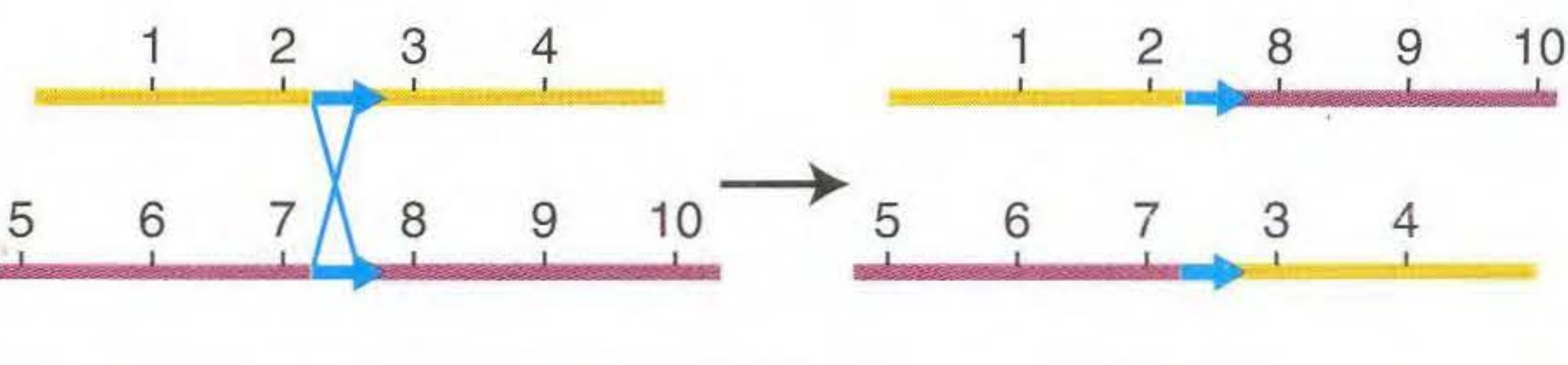

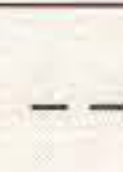


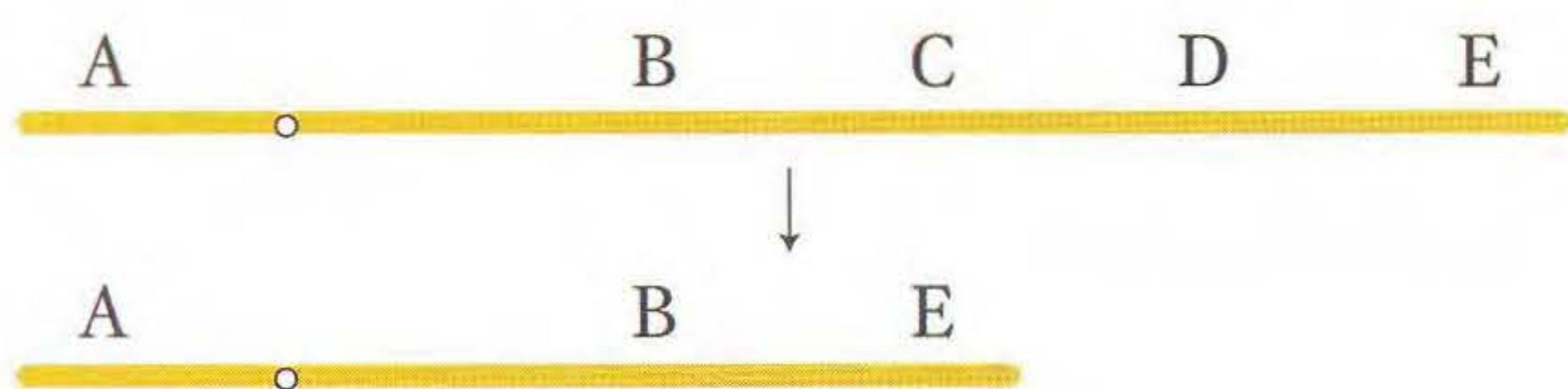
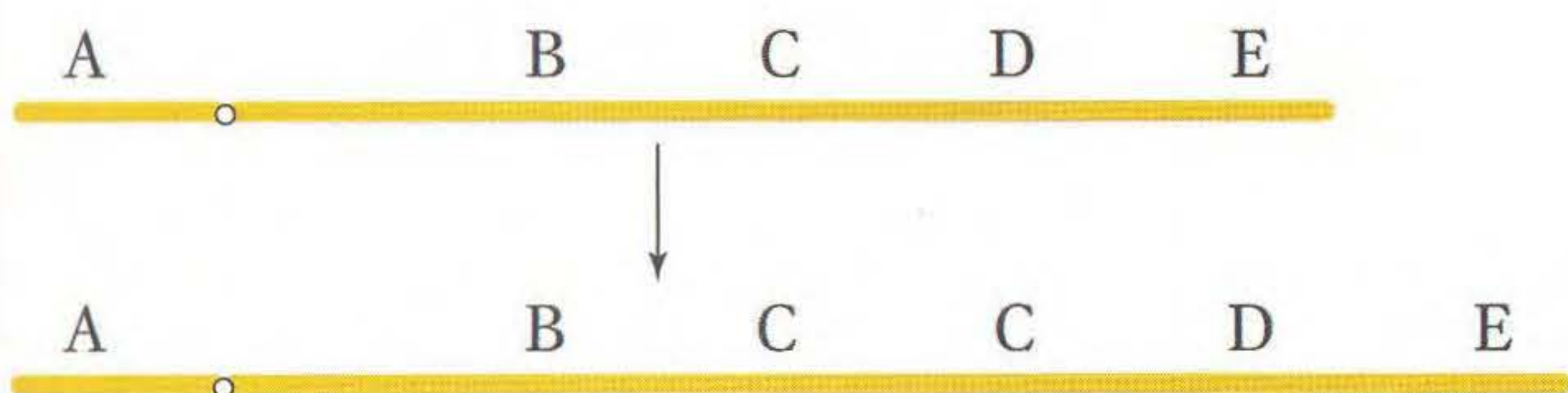
	Quebra e reunião	Crossing entre DNA repetitivo
Deleção		
Deleção e duplicação		
Inversão		
Translocação		
	 Quebra cromossômica  Junção de pontas quebradas  Segmentos de DNA repetitivo  Crossing	

FIG. 16.19 Cada um dos quatro tipos de rearranjos cromossômicos pode ser produzido por um dos dois mecanismos básicos: quebra cromossômica e reunião ou crossing over entre DNA repetitivo. As regiões cromossômicas são numeradas de 1 a 10. Os cromossomos homólogos são da mesma cor.

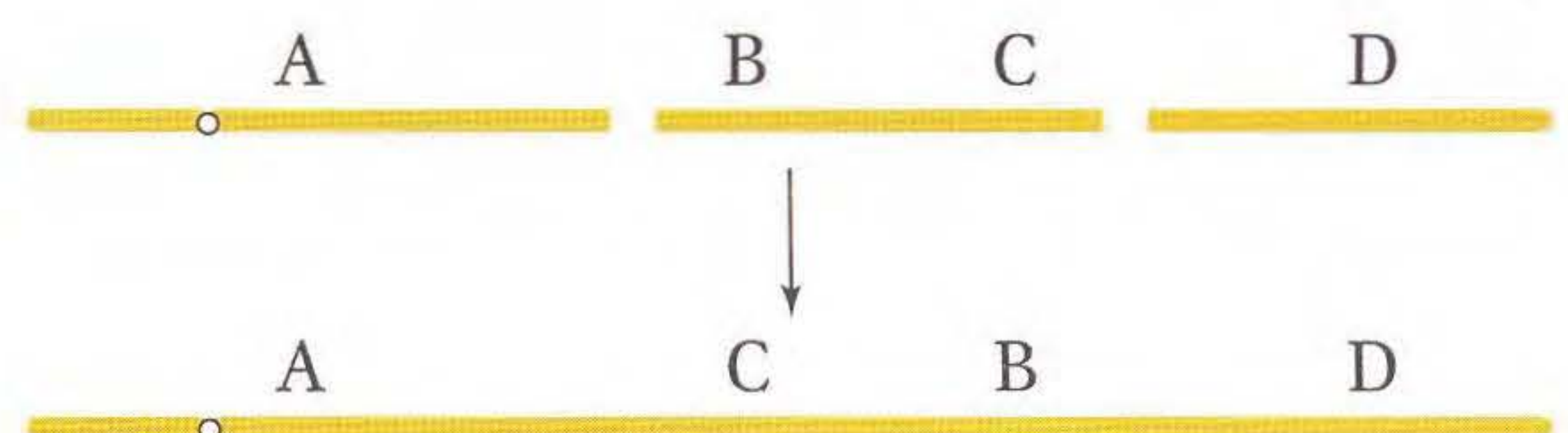


Uma *duplicação* é a repetição de um braço cromossômico. No tipo simples de duplicação, os dois segmentos são adjacentes um ao outro (duplicação em tandem), como na duplicação do segmento C:

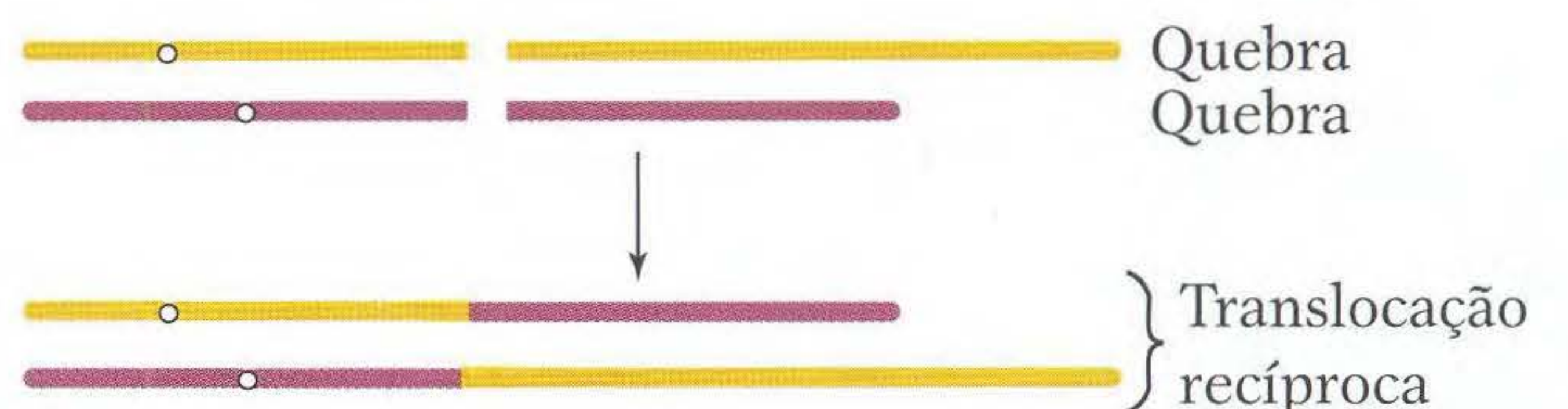


Entretanto, o segmento duplicado pode terminar em uma posição diferente no mesmo cromossomo, ou mesmo em um cromossomo diferente.

Os **rearranjos balanceados** trocam a ordem de genes no cromossomo, mas não removem nem duplicam nenhum DNA. As duas classes simples de rearranjos balanceados são inversões e translocações recíprocas. Uma *inversão* é um rearranjo no qual um segmento interno de um cromossomo foi quebrado duas vezes, girado 180° e reunido.



Uma *translocação recíproca* é um rearranjo no qual dois cromossomos não-homólogos são quebrados uma vez, criando fragmentos acêntricos, que então trocam de lugar:



Às vezes, as quebras de DNA que precedem a formação de um rearranjo ocorrem *dentro* dos genes. Quando ocorrem, elas perturbam o funcionamento do gene porque parte do gene move-se para um novo local e não pode ser feito um transcrito completo. Além disso, as seqüências de DNA em uma das pontas reunidas de um cromossomo rearranjado são seqüências que não são normalmente justapostas. Às vezes, a junção ocorre de tal modo que a fusão produz um gene híbrido não-funcional composto de partes de outros genes.

As seções seguintes consideram as propriedades desses rearranjos balanceados e desbalanceados.

Deleções

Uma deleção é simplesmente a perda de uma parte de um braço cromossômico. O processo de deleção requer duas quebras cromossômicas para remover o segmento intercalar. O fragmento deletado não tem centrômero; conseqüentemente, não pode ser levado para um pólo do fuso na divisão celular e é perdido. O efeito das deleções depende de seu tamanho. Uma pequena deleção *dentro* de um gene, chamada de **deleção intragênica**, inativa o gene e tem o mesmo efeito que outras mutações nulas desse gene. Se o fenótipo homozigoto nulo é viável (como, por exemplo, no albinismo humano), a deleção homozigota também será viável. As deleções intragênicas podem ser distintas das mutações causadas por mudanças de um só nucleotídeo porque os genes com tais deleções nunca reverterem ao tipo selvagem.

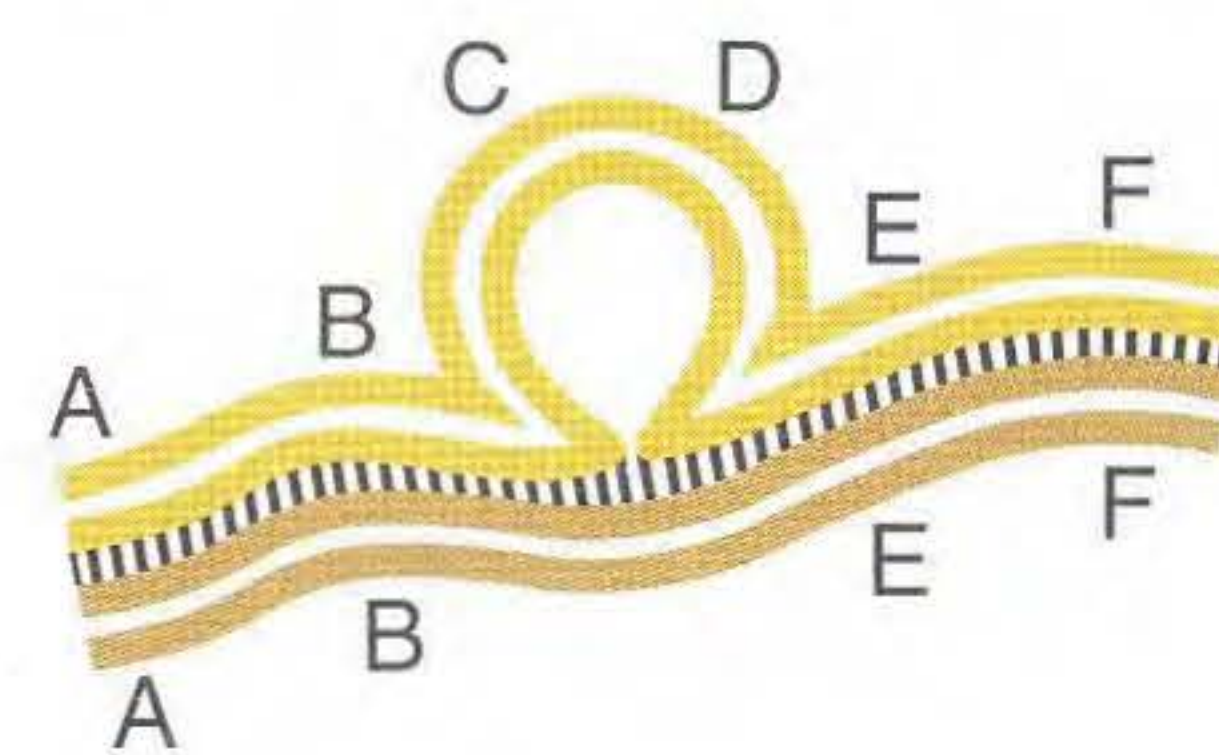
Na maioria desta seção, estaremos lidando com **deleções multigênicas**, nas quais alguns ou muitos genes estão faltando. As conseqüências dessas deleções são mais graves que as de deleções intragênicas. Se tal deleção se torna homozigota por endogamia (isto é, se ambos os homólogos têm a mesma deleção), a combinação é sempre letal. Esse fato sugere que todas as regiões dos cromossomos são essenciais para a viabilidade normal, e a eliminação completa de qualquer segmento do genoma é deletéria. Mesmo um organismo individual heterozigoto para uma deleção multigênica, isto é, tendo um homólogo normal e um que leva à deleção, pode não sobreviver. Principalmente, esse resultado letal é devido à perturbação do balanço gênico normal. Alternativamente, a deleção pode revelar alelos recessivos deletérios, permitindo que cópias únicas sejam expressas.

A letalidade de grandes deleções heterozigotas pode ser explicada pelo desequilíbrio gênico e expressão de recessivos deletérios.

Pequenas deleções às vezes são viáveis em combinação com um homólogo normal. Tais deleções podem ser identificadas examinando-se cromossomos meióticos ao microscópio. A falha do segmento correspondente no cromossomo homólogo em parear cria uma visível **alça de deleção** (Figura 16.20a). Em *Drosophila*, as alças de deleção também são visíveis nos **cromossomos politênicos**. Esses cromossomos são encontrados nas células de glândulas salivares e outros tecidos específicos de alguns insetos. Nessas células, os homólogos fazem pares e replicam-se muitas vezes, e, assim, cada cromossomo é representado por um feixe espesso de réplicas. Esses cromossomos politênicos são facilmente visíveis, e cada um tem um conjunto de bandas de cor escura de posições e número fixo. Essas bandas atuam como marcos cromossômicos úteis. Um exemplo de um cromossomo politênico no qual um homólogo original leva uma deleção é mostrado na Figura 16.20b. Uma deleção pode ser atribuída a um local cromossômico específico examinando-se cromossomos poli-

Alças de deleção em *Drosophila*

(a) Cromossomos meióticos



(b) Cromossomos politênicos

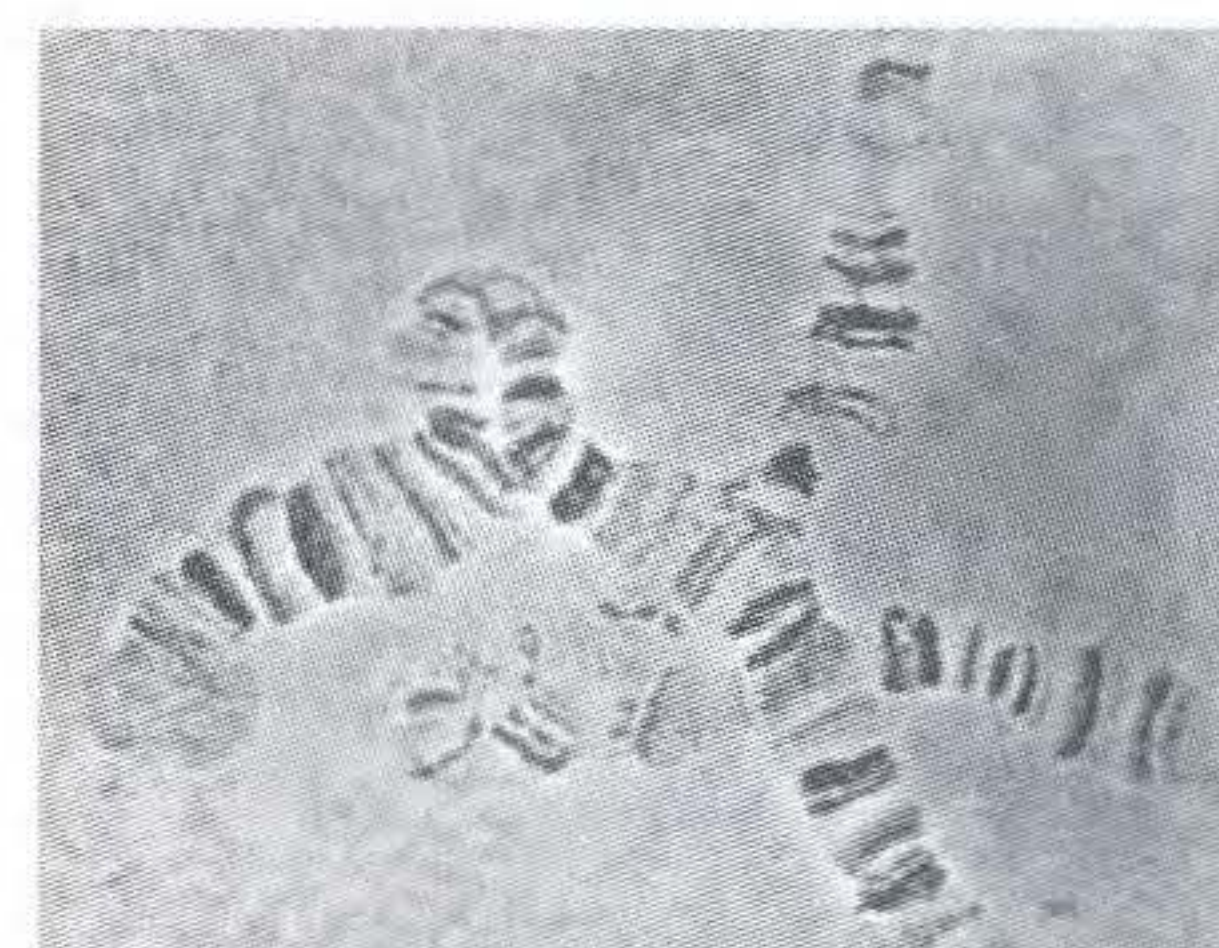
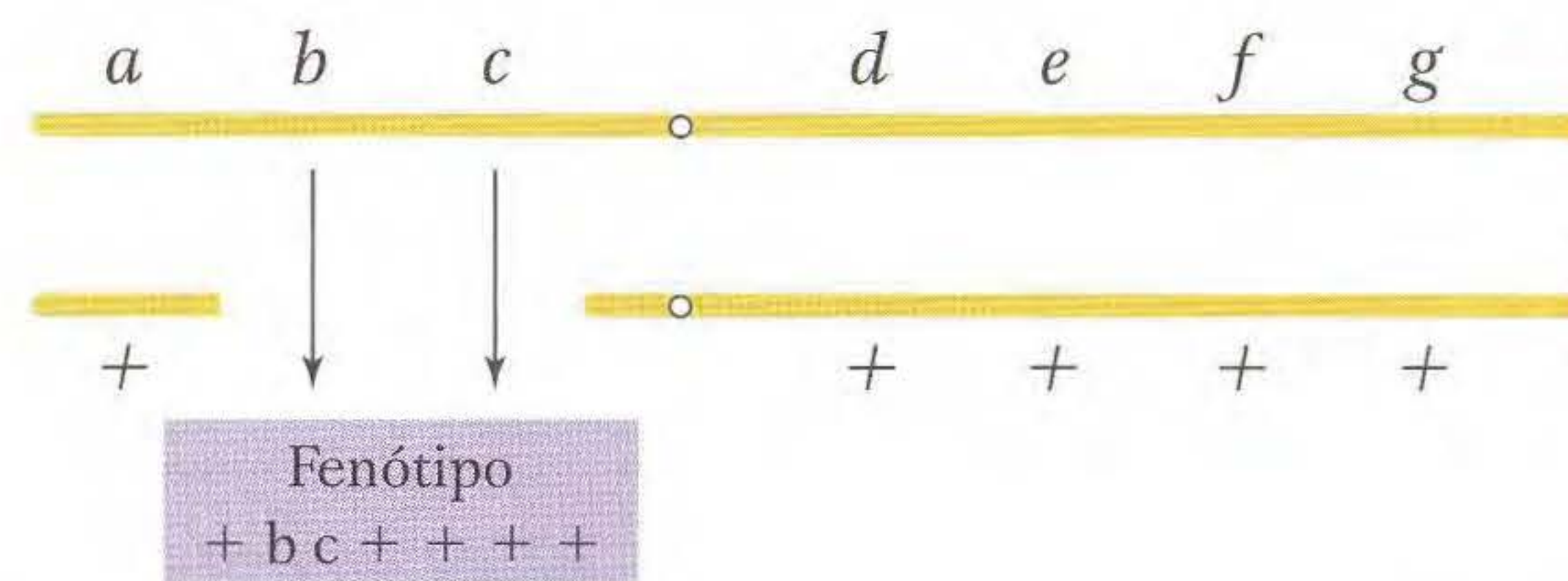


FIG. 16.20 Na meiose, os cromossomos de uma deleção heterozigota de uma configuração em alça. (a) No pareamento meiótico, o homólogo normal forma uma alça. Os genes nessa alça não têm alelos com os quais fazer sinapse. (b) Como os cromossomos politênicos de *Drosophila* (encontrados nas glândulas salivares e outros locais específicos) têm padrões de bandeamento específicos, podemos deduzir quais bandas estão faltando no homólogo com a deleção observando quais bandas aparecem na alça do homólogo normal. [(b) De William M. Gelbart.]

tênicos microscopicamente e determinando a posição da alça de deleção.

Outro indício da presença de uma deleção é que a deleção de um segmento em um homólogo às vezes revela alelos recessivos presentes no outro homólogo, levando à sua inesperada expressão. Considere, por exemplo, a deleção mostrada no seguinte diagrama:



Se não há deleção, não se espera que nenhum dos sete alelos recessivos seja expresso. Entretanto, se *b* e *c* são expressos, então uma deleção envolvendo os genes *b*⁺ e *c*⁺ provavelmente ocorreu no outro homólogo. Como os alelos recessivos parecem mostrar dominância em tais casos, o efeito é chamado de **pseudodominância**.

No caso reverso, se já sabemos a localização da deleção, podemos aplicar o efeito de pseudodominância no sentido oposto para mapear as posições de alelos mutantes. Esse procedimento, chamado de **mapeamento de deleção**, parecia mutações com um conjunto de deleções superpostas definidas. Um exemplo de *Drosophila* é mostrado na Figura 16.21. Nesse diagrama, o mapa de recombinação é mostrado

Origem da deleção da síndrome do *cri du chat*

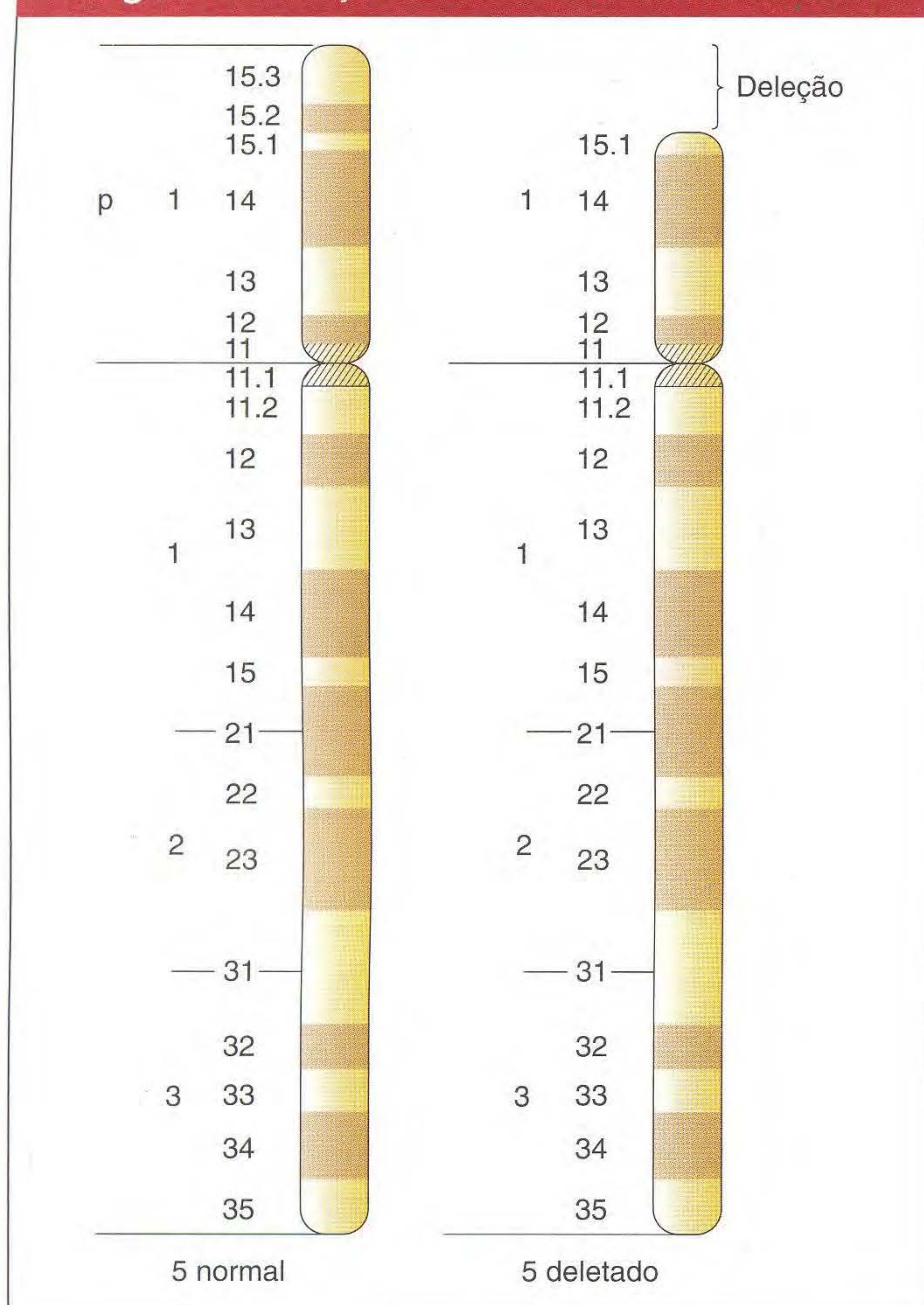


FIG. 16.22 A síndrome do *cri du chat* é causada pela perda da ponta do braço curto de um dos homólogos do cromossomo 5.

A maioria das deleções humanas, tais como as que consideramos, surgem espontaneamente nas gônadas de um genitor normal de uma pessoa afetada; assim, geralmente não são encontrados sinais das deleções nos cromossomos dos genitores. Menos comumente, os indivíduos portadores de deleções aparecem na prole de um indivíduo contendo um rearranjo balanceado de cromossomos não-detectado. Por exemplo, a síndrome do *cri du chat* pode resultar de um genitor heterozigoto para uma translocação recíproca, pois (como veremos) a segregação produz deleções. As deleções também podem resultar de recombinação dentro dos heterozigotos tendo uma inversão pericêntrica (uma inversão que envolve o centrômero) em um cromossomo. Ambos os mecanismos serão detalhados mais adiante neste capítulo.

Os animais e plantas apresentam diferenças na sobrevivência de gametas ou prole que possua deleções. Um animal masculino com uma deleção em um cromossomo produz espermatozoides levando um ou outro dos dois cromossomos em números aproximadamente iguais. Esses espermatozoides parecem funcionar até certo ponto independentemente de seu conteúdo genético. Nas plantas diplóides, por outro lado, o pólen produzido por uma deleção heterozigota é de dois tipos: pólen fun-

Provável origem da deleção da síndrome de Williams

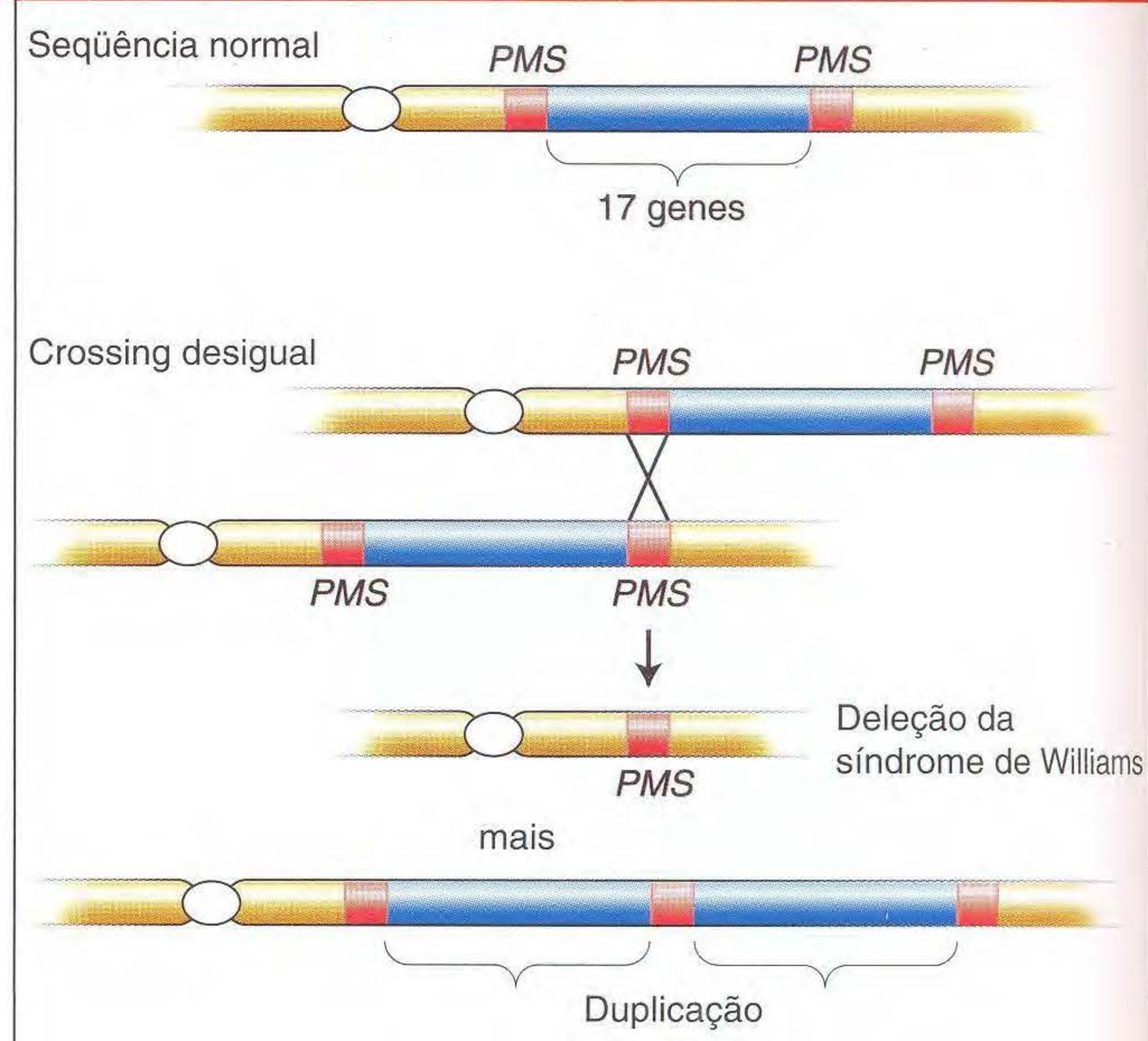


FIG. 16.23 Um crossing entre genes flangeadores repetitivos esquerdo e direito resulta em rearranjos recíprocos, um dos quais corresponde à deleção da síndrome de Williams.

cional, levando o cromossomo normal, e pólen não-funcional (abortado), levando o homólogo deficiente. Assim, as células de pólen parecem ser sensíveis a mudanças na quantidade de material cromossômico, e essa sensibilidade pode agir para eliminar as deleções. Esse efeito é análogo à sensibilidade do pólen à aneuploidia de cromossomo inteiro, já descrita neste capítulo. Ao contrário das células de espermatozoides animais, cuja atividade metabólica depende de enzimas que já foram depositadas nelas durante sua formação, as células do pólen devem germinar e, então, produzir um longo tubo polínico que cresce para fertilizar o ovócito. Esse crescimento requer que a célula do pólen produza grandes quantidades de proteína, tornando-o sensível a anomalias genéticas em seu próprio núcleo. Os ovócitos das plantas, em contraste, são bem tolerantes a deleções, supostamente porque recebem sua nutrição dos tecidos maternos circundantes.

Duplicações

Os processos de mutação cromossômica às vezes produzem uma cópia extra de alguma região cromossômica. As regiões duplicadas podem ser situadas adjacentes uma à outra, chamada de **duplicação em tandem**, ou a cópia extra pode estar situada em outra parte do genoma, chamada de **duplicação insercional**. Uma célula diplóide contendo uma duplicação terá três cópias da região cromossômica em questão: duas em um conjunto cromossômico e uma no outro, um exemplo de duplicação heterozigota. Na prófase meiótica, os heterozigotos para duplicação em tandem mostram uma alça consistindo na região extra não-pareada.

As duplicações sintéticas de cobertura conhecida podem ser usadas para mapeamento genético. Nos haplóides, por

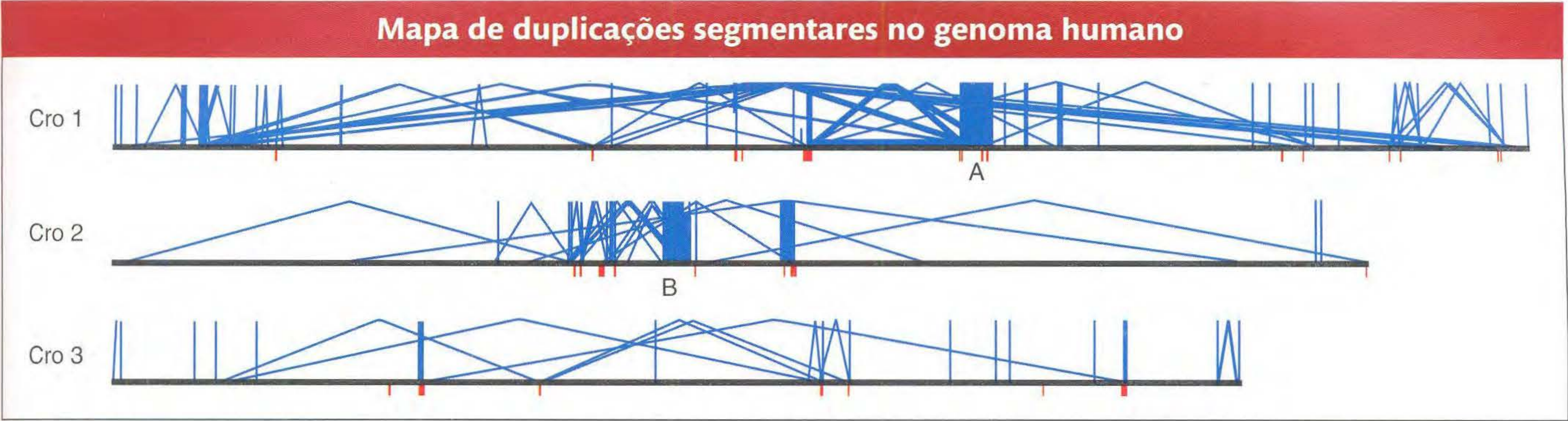


FIG. 16.24 O mapa dos cromossomos humanos 1, 2 e 3 mostra as posições das duplicações com mais de 10 quilobases de tamanho. As linhas azuis de ligação mostram duplicações intracromossômicas (a grande maioria). As duplicações intercromossômicas são mostradas com barras vermelhas. As letras A e B indicam pontos quentes onde as recombinações e duplicações deram origem a distúrbios genéticos. [De J.A. Bailey et al., "Recent Segmental Duplications in the Human Genome," *Science* 297, 2002, 1003–1007.]

exemplo, uma linhagem cromossomicamente normal levando uma nova mutação recessiva *m* pode ser cruzada com linhagens tendo um número de rearranjos geradores de duplicação (p. ex., translocações e inversões pericêntricas). Em qualquer cruzamento, se alguma prole de duplicação tem o fenótipo recessivo, a duplicação não envolve o gene *m*, pois, se o fizer, seu segmento extra iria mascarar o alelo recessivo *m*.

As análises de seqüências de DNA genômico revelaram um alto nível de duplicações em humanos e na maioria dos organismos-modelo. Seqüências simples repetidas, que são amplas no genoma e úteis como marcadores moleculares no mapeamento, foram discutidas em capítulos anteriores. Entretanto, outra classe de duplicações baseadas em unidades duplicadas é muito maior do que seqüências simples repetidas. As duplicações nessa classe são chamadas de **duplicações segmentares**. As unidades duplicadas em duplicações segmentares variam de 10 a 50 quilobases de tamanho e englobam genes inteiros e regiões entre eles. O tamanho das duplicações segmentares é mostrado na Figura 16.24, na qual a maioria das duplicações são dispersas, mas existem alguns casos de tandem. Outra propriedade mostrada na Figura 16.24 é que a dispersão das unidades duplicadas é principalmente dentro do mesmo cromossomo, e não entre os cromossomos. A origem das duplicações segmentares ainda é desconhecida.

As duplicações segmentares são tidas como tendo um papel importante como substratos para recombinação homóloga não-alélica, como mostrado na Figura 16.19. O crossing over entre duplicações segmentares pode levar a vários rearranjos cromossômicos. Esses rearranjos parecem ter tido importância na evolução, pois algumas grandes inversões que são importantes diferenças entre seqüências de humanos e macacos quase certamente vieram de NAHR. Também parece provável que NAHR tenha sido responsável por rearranjos que causam algumas doenças humanas. Os loci de tais doenças são pontos quentes de duplicação segmentar. Os exemplos de tais loci são mostrados na Figura 16.24.

Já vimos que, em alguns organismos, tais como poliplóides, o genoma atual evoluiu como resultado de uma duplicação

de todo um genoma ancestral. Quando ocorre a duplicação de um genoma inteiro, cada gene é duplicado. Esses genes duplicados são uma fonte de algumas das duplicações segmentares encontradas nos genomas. Um caso bem estudado é o da levedura do pão, *Saccharomyces cerevisiae*. A evolução desse genoma foi analisada comparando a seqüência do genoma inteiro de *S. cerevisiae* com o de outra levedura, *Kluyveromyces*, cujo genoma é similar ao genoma ancestral de levedura. Aparentemente, no curso da evolução de *Saccharomyces*, o genoma ancestral similar ao de *Kluyveromyces* dobrou, e, assim, existem dois conjuntos, cada um contendo o genoma inteiro. Após ocorrer a duplicação, muitas cópias do gene foram perdidas de um conjunto ou do outro, e os restantes foram rearranjados, resultando no genoma atual de *Saccharomyces*. Esse processo é reconstruído na Figura 16.25.

Inversões

Já vimos que, para criar uma inversão, um segmento de um cromossomo é cortado, girado e reinserido. As inversões são de dois tipos básicos. Se o centrômero fica fora da inversão, ela é dita **paracêntrica**. As inversões que envolvem o centrômero são **pericêntricas**.

	A	B		C	D	E	F
Seqüência normal							
Paracêntrica	A	B		C	E	D	F
Pericêntrica	A	D	C		B	E	F

Como as inversões são rearranjos balanceados, elas não alteram a quantidade geral de material genético, e, assim, não resultam em desequilíbrio gênico. Os indivíduos com inversões em geral são normais, se não ocorrerem quebras dentro dos genes. Uma quebra que perturbe um gene produz uma mutação que pode ser detectável como um fenótipo anormal. Se o gene tem uma função essencial, então o ponto de quebra atua como uma mutação letal ligada à inversão. Em tal caso, a inversão não pode ser levada à homozigose. Entretanto, muitas inversões podem tornar-se homozigotas,

Surgem duplicações no genoma de *Saccharomyces* por poliploidia ancestral

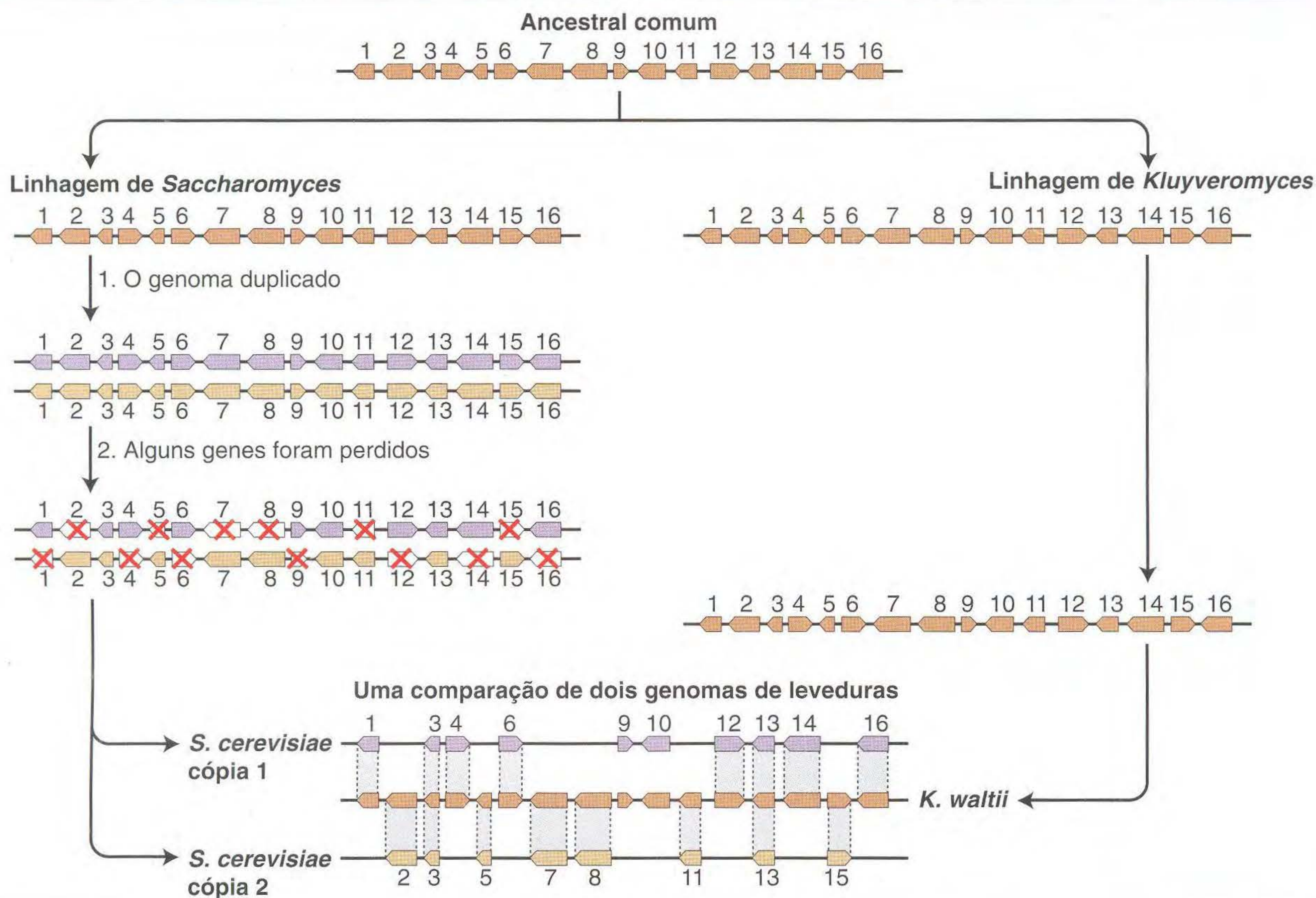


FIG. 16.25 Um ancestral comum similar à moderna levedura *Kluyveromyces* duplicou seu genoma (1). Alguns genes foram perdidos (2). Os genes duplicados, tais como 3 e 13, estão na mesma ordem relativa. O painel inferior compara os dois genomas modernos. [Adaptado da Figura 1 de Manolis Kellis, Bruce W. Birren, and Eric S. Lander, "Proof and Evolutionary Analysis of Ancient Genome Duplication in the Yeast *Saccharomyces cerevisiae*," *Nature*, vol. 428, April 8, 2004, copyright Nature Publishing Group.]

e, além disso, as inversões podem ser detectadas em organismos haplóides. Nesses casos, os pontos de quebra da inversão claramente não ocorrem em regiões essenciais. Algumas das possíveis consequências da inversão a nível de DNA são mostradas na Figura 16.26.

A maioria das inversões ocorre em células diplóides que contêm um conjunto cromossômico normal mais um conjunto com a inversão. Esse tipo de célula é chamada de **inversão heterozigota**, mas note que essa designação não significa que qualquer locus gênico seja heterozigoto. Significa que estão presentes um conjunto cromossômico normal e um anormal. A localização do segmento invertido em geral pode ser detectada microscopicamente. Na meiose, um cromossomo gira nas extremidades da inversão para parear com seu homólogo não-girado. Desse modo, os homólogos pareados formam uma visível **alça de inversão** (Figura 16.27).

Em uma inversão **paracêntrica**, o crossing dentro da alça de inversão na meiose conecta centrômeros homólogos em uma **ponte dicêntrica** que também produz um **fragmento acêntrico** (Figura 16.28). Então, à medida que os cromossomos se separam na anáfase I, os centrômeros permanecem ligados pela ponte. O fragmento acêntrico não pode alinhar-se nem mover-se. Conseqüentemente, ele é perdido. A tensão acaba rompendo a ponte dicêntrica, formando dois cromossomos

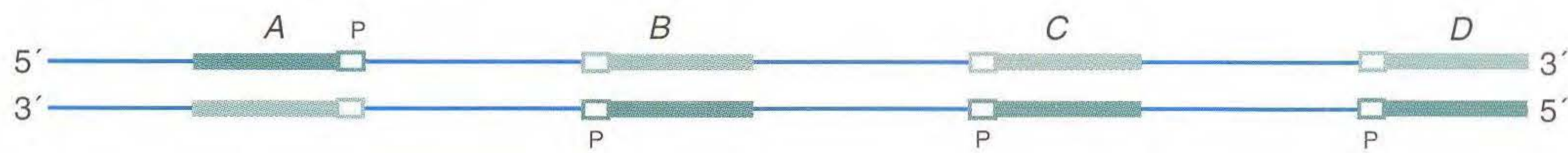
com deleções terminais. Os gametas contendo tais cromossomos ou os zigotos que eles acabem originando provavelmente serão inviáveis. Assim, um evento de crossing, que normalmente gera a classe recombinante de produtos meióticos, é letal para esses produtos. O resultado geral é uma frequência drasticamente menor de recombinantes viáveis. De fato, para genes dentro da inversão, a frequência de recombinantes é próxima de zero. (Não é exatamente zero, porque crossings duplos apenas entre duas cromátides são viáveis.) Para genes flangeadores da inversão, a FR é reduzida na proporção do tamanho da inversão porque, para uma inversão longa, há uma probabilidade maior de ocorrer um crossing dentro dela e produzir um produto meiótico inviável.

Em uma inversão **pericêntrica** heterozigota, o efeito genético final é o mesmo que o de uma inversão paracêntrica; os produtos do crossing não são recuperados, mas os motivos são diferentes. Em uma inversão pericêntrica, os centrômeros estão contidos dentro da região invertida. Conseqüentemente, os cromossomos envolvidos no crossing over separam-se do modo normal, sem a criação de uma ponte (Figura 16.29). Entretanto, o crossing produz cromátides que contêm uma duplicação e uma deleção para partes diferentes do cromossomo. Nesse caso, se um gameta portador de crossing é fertilizado, o zigoto morre devido a desequilíbrio gênico.

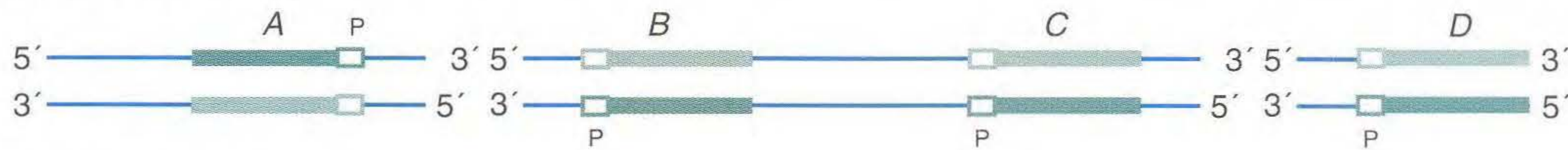
As inversões podem causar uma variedade de mudanças estruturais no DNA

Pontos de quebra *entre* os genes

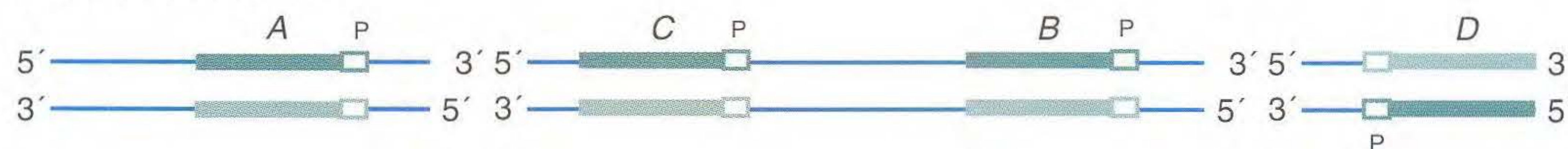
Seqüência normal



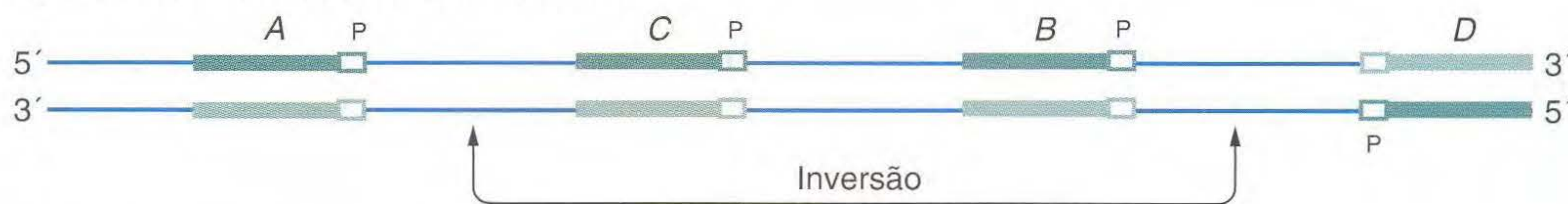
Quebras no DNA



Alinhamento invertido

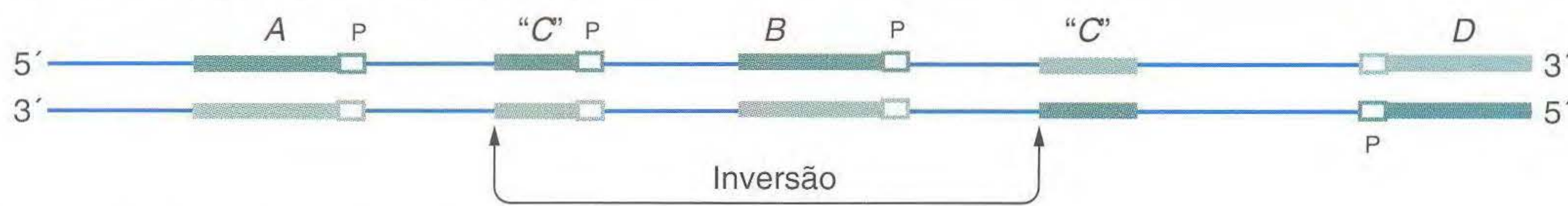


União de quebras para completar inversão



Um ponto de quebra *entre* os genes

Um *dentro* do gene C (C perturbado)



Pontos de quebra *nos* genes A e D

Criando fusões de genes

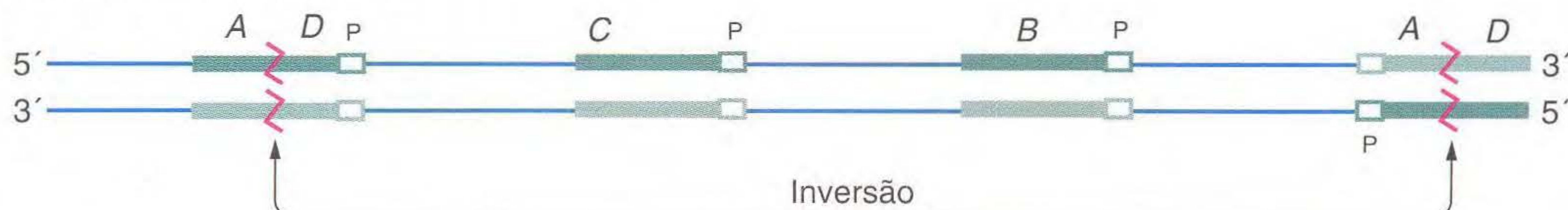


FIG. 16.26 Uma inversão pode não ter efeito nos genes, pode perturbar um gene ou pode fundir partes de dois genes, dependendo do local do ponto de quebra. Os genes são representados por A, B, C e D. O filamento-molde é verde-escuro; o filamento não-molde é verde-claro; as linhas vermelhas onduladas indicam onde as quebras no DNA produzem fusões gênicas (A com D) após inversão e reunião. A letra P indica promotor; as setas indicam as posições dos pontos de quebra.

Novamente, o resultado é que apenas as cromátides não-crossing estão presentes na prole viável. Assim, o valor da FR dos genes dentro de uma inversão pericêntrica também é zero.

As inversões afetam a recombinação também de outro modo. As inversões heterozigotas geralmente têm problemas mecânicos de pareamento na região da inversão. A alça de inversão causa uma grande distorção que pode ir além da própria alça. Essa distorção reduz a oportunidade de crossing over nas regiões vizinhas.

Consideremos um exemplo dos efeitos de uma inversão na frequência de recombinação. Uma *Drosophila* tipo selvagem de uma população natural é cruzada com um estoque homozigoto recessivo de laboratório *dp cn/dp cn*. (O alelo *dp* codifica asas dumpy e *cn* codifica olhos cinnabar. Os dois genes são conhecidos como distantes 45 unidades de mapa no

cromossomo 2.) A geração F_1 é tipo selvagem. Quando uma fêmea F_1 é cruzada com um genitor recessivo, a prole é

250	tipo selvagem	$+/+ dp cn$
246	dumpy cinnabar	$dp cn/dp cn$
5	dumpy	$dp +/dp cn$
7	cinnabar	$+ cn/dp cn$

Nesse cruzamento, que é efetivamente um cruzamento-teste diíbrido, espera-se que 45% da prole sejam dumpy ou cinnabar (elas constituem as classes crossing), mas apenas 12 de 508, cerca de 2%, são obtidas. Algo está reduzindo o crossing over nessa região, e a explicação provável é uma inversão envolvendo a maior parte da região *dp-cn*. Como a FR esperada foi baseada em medidas feitas em linhagens laboratoriais, a mosca tipo selvagem da natureza foi mais provavelmente a

Alças de inversão na meiose

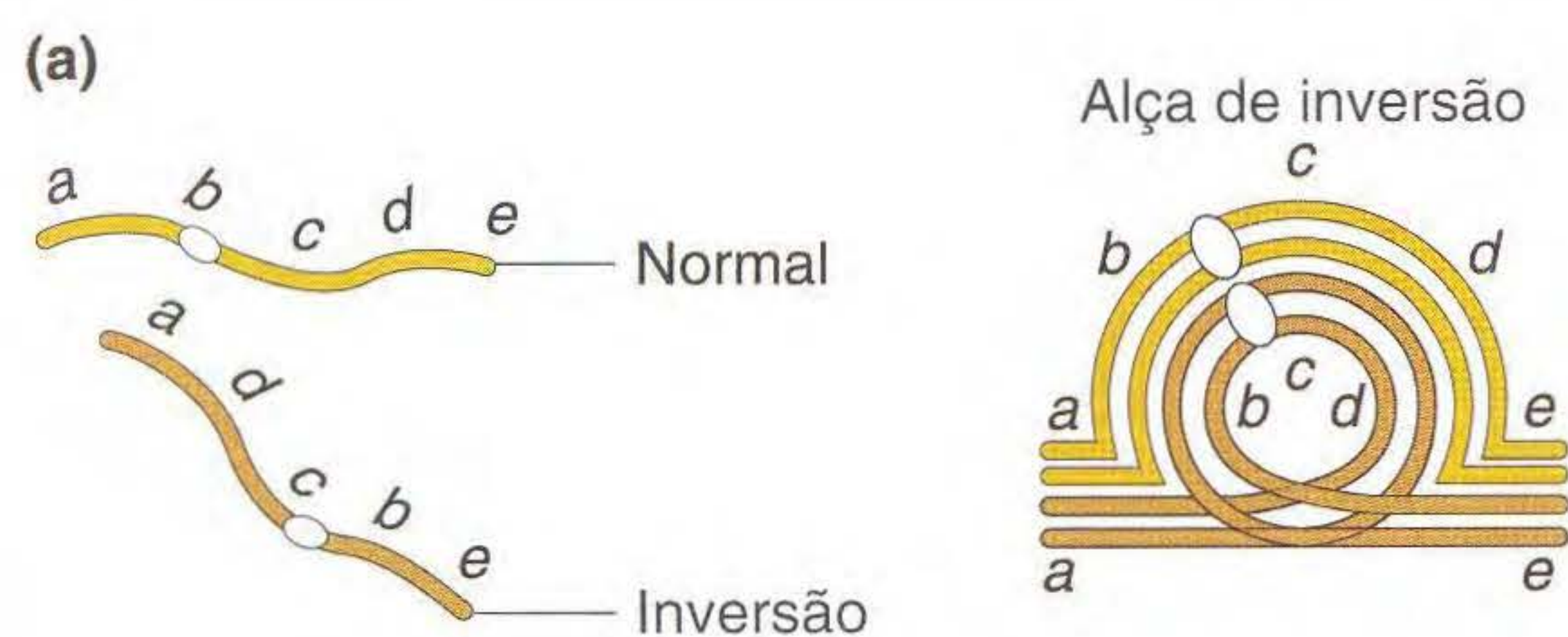
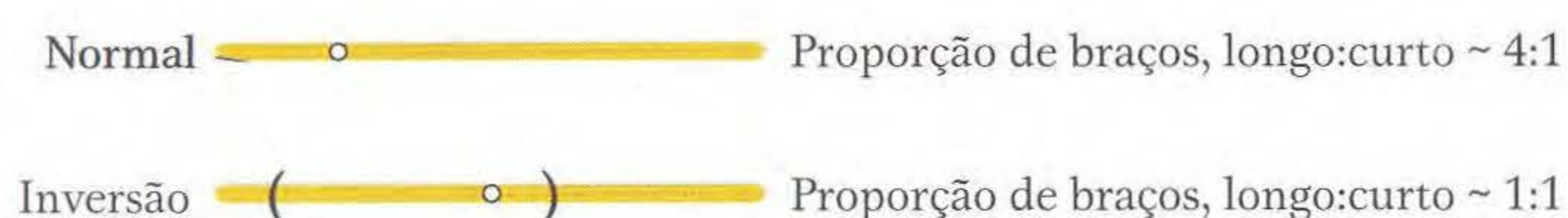


FIG. 16.27 Os cromossomos de inversões heterozigotas formam par em uma alça na meiose. (a) Representação diagramática. (b) Micrografia eletrônica de complexos sinaptonêmicos na prófase I da meiose em um camundongo heterozigoto para uma inversão paracêntrica. São mostrados três meiócitos diferentes. [(b) De M.J. Moses, Department of Anatomy, Duke Medical Center.]

fonte do cromossomo invertido. Assim, o cromossomo 2 na F_1 pode ser representado do seguinte modo:



As inversões pericêntricas também podem ser detectadas microscopicamente por novas proporções de braços. Considere a seguinte inversão pericêntrica:



Note que a proporção de tamanho do braço longo para o braço curto foi alterada de cerca 4:1 para cerca de 1:1 pela inversão. As inversões paracêntricas não alteram a proporção dos braços, mas podem ser detectadas microscopicamente observando-se as mudanças no bandejamento ou outros marcadores cromossômicos, se disponíveis.

As inversões paracêntricas podem levar a produtos de deleção

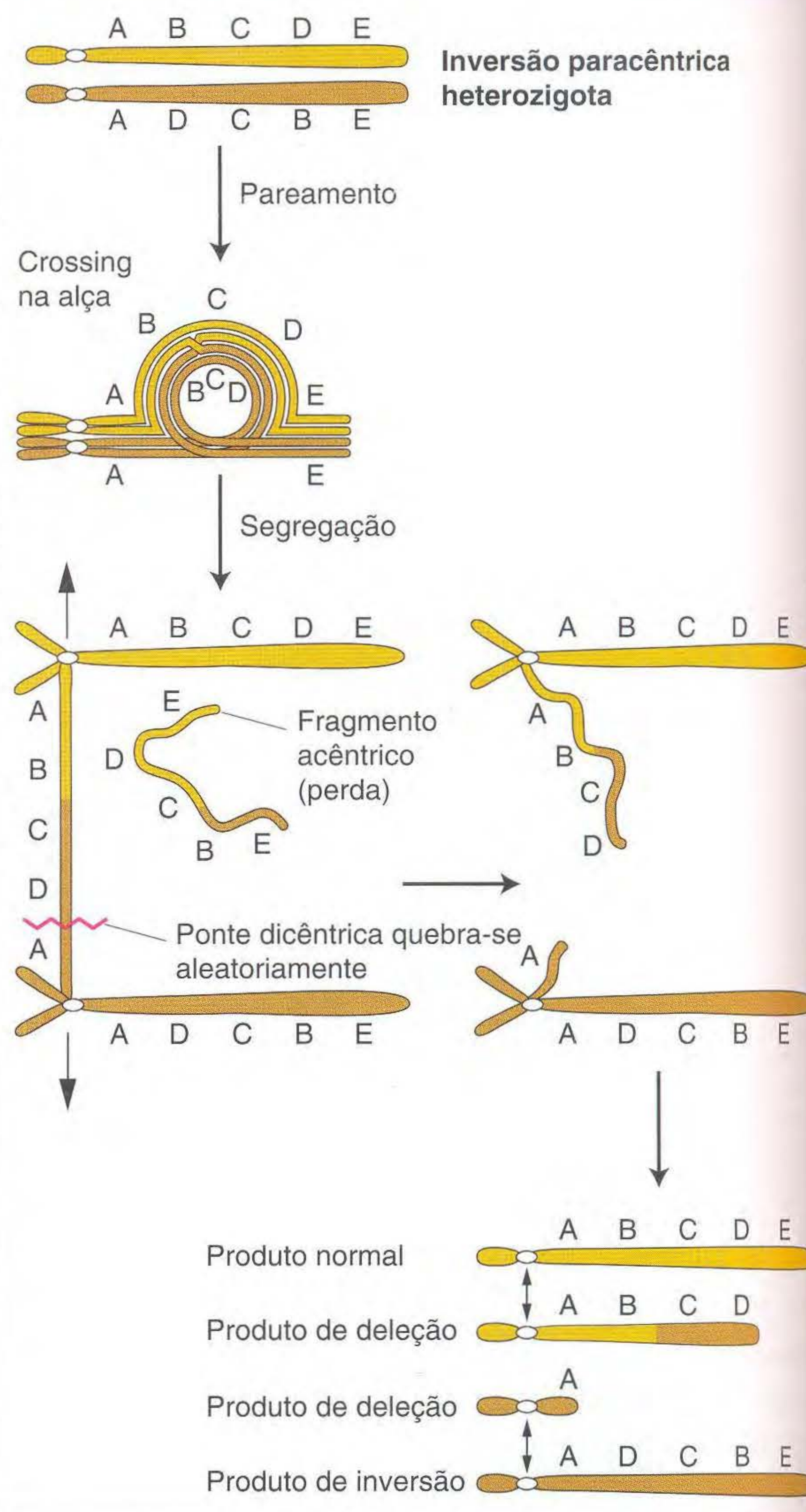


FIG. 16.28 Um crossing na alça de um heterozigoto para inversão paracêntrica dá origem a cromossomos contendo deleções.

As principais características diagnósticas das inversões heterozigotas são as alças de inversão, frequência de recombinantes reduzida e fertilidade reduzida devido a produtos meióticos deletados ou desbalanceados.

Em alguns sistemas-modelo experimentais, notadamente *Drosophila* e o nematódeo *Caenorhabditis elegans*, as inversões são usadas como balanceadores. Um cromossomo **balanceador** contém *múltiplas* inversões; assim, quando é combinado ao cromossomo tipo selvagem correspondente, não pode haver produtos crossing viáveis. Em algumas análises, é importante ter estoque com todos os alelos em um cromossomo juntos. O geneticista cria indivíduos tendo genomas que combinam tal cromossomo com um balanceador. Essa combinação elimina crossings, e, assim, apenas as

As inversões pericêntricas podem levar a produtos de duplicação e deleção

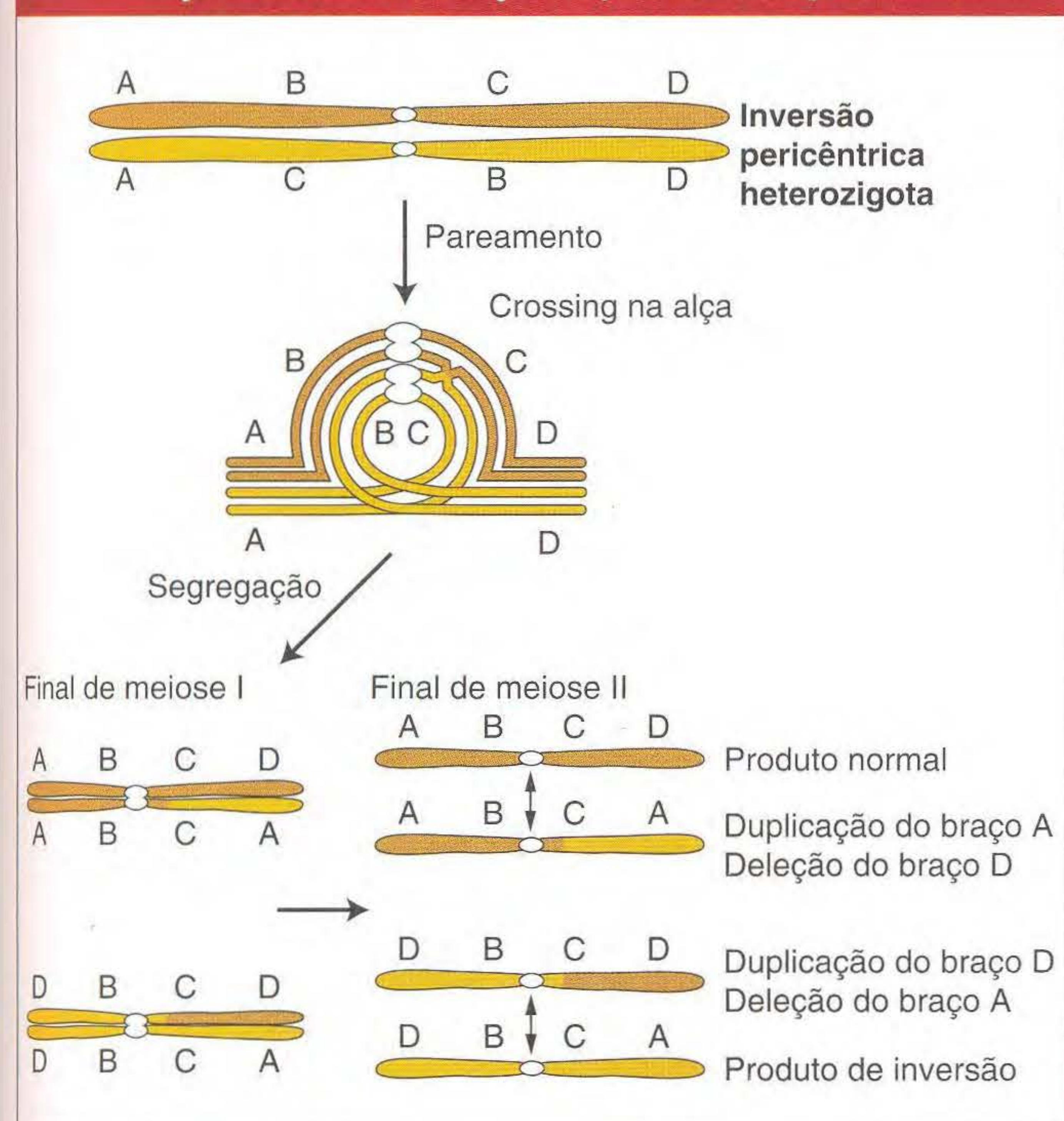


FIG. 16.29 Um crossing na alça de um heterozigoto para inversão pericêntrica origina cromossomos contendo duplicações e deleções.

combinações parentais aparecem na prole. Por conveniência, os cromossomos balanceadores são marcados com uma mutação morfológica dominante. O marcador permite que o geneticista siga a segregação de todo o balanceador ou seu homólogo normal notando a ausência ou presença do marcador.

Translocações recíprocas

Existem vários tipos de translocações, mas aqui consideraremos apenas as translocações recíprocas, o tipo mais simples. Lembre que, para formar uma translocação recíproca, dois cromossomos trocam fragmentos acêntricos criados por duas quebras cromossômicas simultâneas. Como em outros rearranjos, a meiose nos heterozigotos tendo dois cromossomos translocados e suas contrapartes normais produz configurações características. A Figura 16.30 ilustra a meiose em um indivíduo que é heterozigoto para uma translocação recíproca. Note a configuração de pareamento em forma de cruz. Devido à ação da lei da distribuição independente, existem dois padrões comuns de segregação. Usemos N_1 e N_2 para representar os cromossomos normais, e T_1 e T_2 para os cromossomos translocados. A segregação de cada um dos cromossomos estruturalmente normais com um dos translocados ($T_1 + N_2$ e $T_2 + N_1$) é chamada de **segregação adjacente-1**. Cada um dos produtos meióticos é deficiente para um braço diferente do cruzamento e tem uma duplicação do outro. Esses produtos são inviáveis. Por outro lado, os dois cromossomos normais podem segregar juntos, como farão as partes recíprocas dos translocados, para gerar os produtos

$N_1 + N_2$ e $T_1 + T_2$. Esse padrão de segregação é chamado de **segregação alternada**. Esses produtos são ambos balanceados e viáveis.

As segregações adjacente-1 e alternada são iguais em número, e, assim, metade da população geral de gametas será não-funcional, uma condição conhecida como **semi-esterilidade** ou “meia-esterilidade”. A semi-esterilidade é uma ferramenta diagnóstica importante para identificar translocações heterozigotas. Entretanto, a semi-esterilidade é definida diferentemente para plantas e animais, e 50% dos produtos meióticos que são de segregação adjacente-1 geralmente abortam no estágio gamético (Figura 16.31). Em animais, esses produtos são viáveis como gametas, mas letais para os zigotos que produzem na fertilização.

Lembre que os heterozigotos para inversões também podem apresentar alguma redução em fertilidade, mas em uma quantidade dependente do tamanho da região afetada. A redução exata de 50% em gametas viáveis ou zigotos é geralmente um indício diagnóstico confiável para uma translocação.

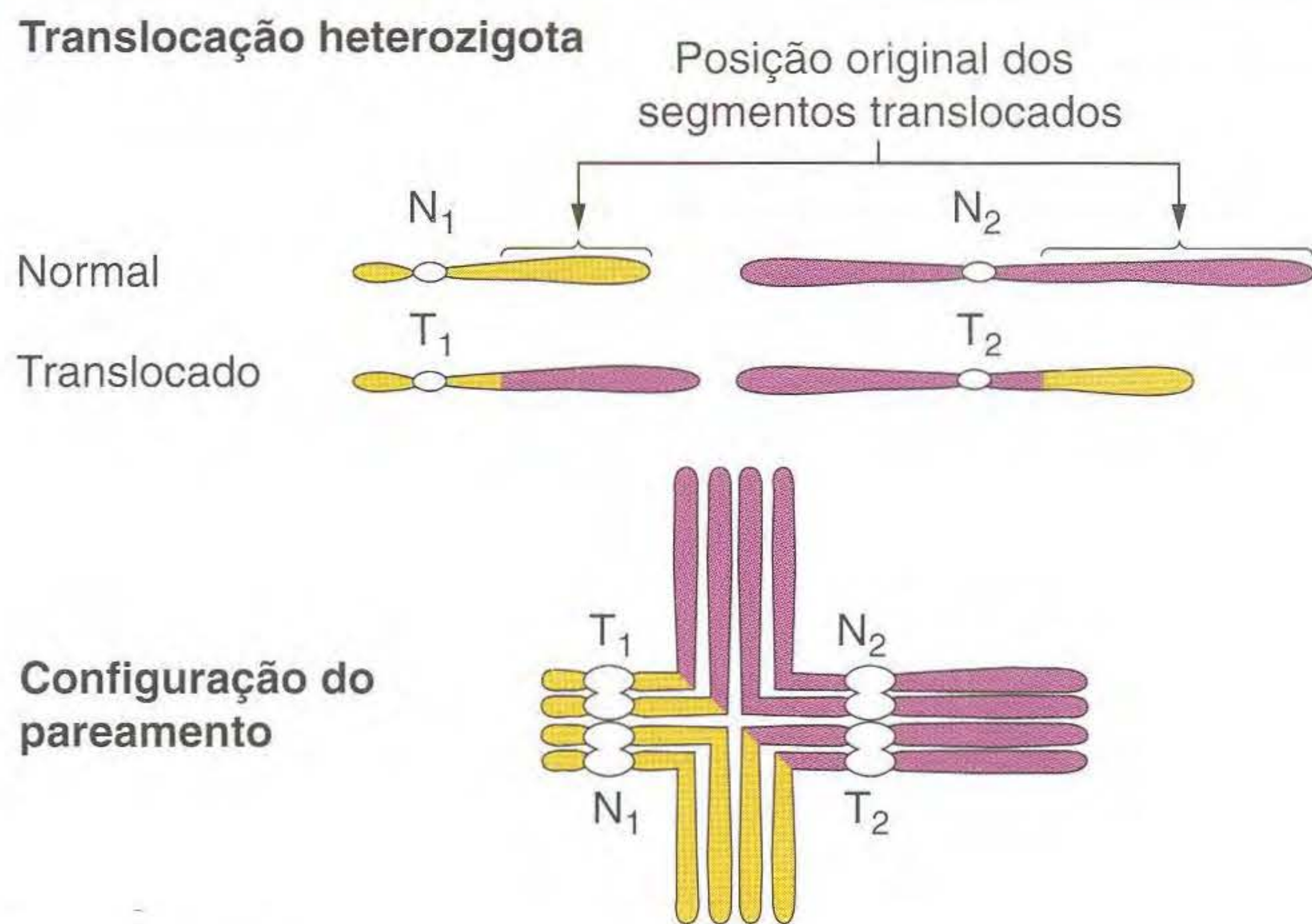
Geneticamente, os genes em cromossomos translocados atuam como se estivessem ligados se seus loci estiverem próximos ao ponto de quebra da translocação. A Figura 16.32 mostra um heterozigoto para a translocação que foi estabelecido cruzando um indivíduo $a/a ; b/b$ com um homozigoto para translocação possuindo alelos do tipo selvagem. Quando o heterozigoto é submetido a cruzamento-teste, os recombinantes são criados, mas não sobrevivem, porque eles levam genomas desbalanceados (duplicações e deleções). A única prole viável são os que possuem genótipos parentais; assim, a ligação é vista entre loci que, originalmente, estavam em cromossomos diferentes. A ligação aparente dos genes normalmente conhecidos como estando em cromossomos não-homólogos separados, às vezes chamada de **pseudoligação**, é um indício diagnóstico genético da presença de uma translocação.

As translocações recíprocas heterozigotas são diagnosticadas geneticamente pela semi-esterilidade e pela aparente ligação de genes cujos loci normais estão em cromossomos separados.

Translocações robertsonianas

Voltemos à família com uma criança portadora da síndrome de Down, apresentada no começo do capítulo. O nascimento pode ser uma coincidência, pois as coincidências acontecem. Entretanto, o aborto dá um indício de que algo mais pode estar ocorrendo. Uma grande proporção de abortos espontâneos leva anomalias cromossômicas; logo, talvez seja esse o caso no exemplo. Considerando que sim, o casal pode ter tido duas concepções com mutações cromossômicas, o que seria muito improvável, a menos que houvesse uma causa comum. Entretanto, uma pequena proporção de casos de síndrome de Down resulta de translocação em um dos genitores. Vimos que as translocações podem produzir uma prole que tem material extra de parte do genoma, e, assim, uma transloca-

Os dois principais padrões de segregação cromossômica em uma translocação recíproca heterozigota



Dois tipos de segregações:

Adjacente-1	Produtos meióticos finais
Para cima $T_1 + N_2$	Duplicação de púrpura, deleção do segmento laranja translocado
Para baixo $N_1 + T_2$	Duplicação de laranja, deleção do segmento translocado púrpura
Alternada	
Para cima $T_1 + T_2$	Genótipo de translocação
Para baixo $N_1 + N_2$	Normal

FIG. 16.30 Os cromossomos segregantes de um heterozigoto para translocação recíproca formam uma configuração de pareamento em cruz. Os dois padrões de segregação mais comumente encontrados que resultam são o geralmente inviável "adjacente-1" e o viável "alternado". N_1 e N_2 cromossomos não-homólogos normais; T_1 e T_2 , cromossomos translocados. Acima e abaixo designam os pólos opostos para os quais migram os homólogos na anáfase I.

ção envolvendo o cromossomo 21 pode produzir uma prole que tem material extra desse cromossomo. Na síndrome de Down, a translocação responsável é do tipo chamado de *translocação robertsoniana*. Ela produz uma prole levando uma cópia extra quase completa do cromossomo 21. A transloca-

ção e sua segregação são ilustradas na Figura 16.33. Note que, além dos complementos que causam a síndrome de Down, outros complementos cromossômicos aberrantes são produzidos, a maioria dos quais abortam. Em nosso exemplo, o homem pode ter essa translocação, que ele pode ter herdado de sua avó. Para confirmar essa possibilidade, são verificados seus cromossomos. Seu filho não-afetado pode ter cromossomos normais ou pode ter herdado sua translocação.



FIG. 16.31 Pólen de uma planta de milho semi-estéril. Os grãos de pólen claros contêm produtos meióticos cromossomicamente desbalanceados de um heterozigoto para translocação recíproca. Os grãos de pólen opacos, que contêm o genótipo completo da translocação ou os cromossomos normais, são funcionais na fertilização e desenvolvimento. [William Sheridan.]

Aplicações de inversões e translocações

As inversões e translocações demonstraram-se ferramentas genéticas úteis; alguns exemplos de seus usos são dados a seguir.

Mapeamento gênico As inversões e translocações são úteis para o mapeamento e isolamento subsequente de genes específicos. O gene para neurofibromatose humana foi isolado desse modo. A informação crítica veio de pessoas que não só tinham a doença, mas também possuíam translocações cromossômicas. Todas as translocações tinham um ponto de quebra em comum, em uma banda próxima ao centrômero do cromossomo 17. Assim, essa banda parecia ser o locus do gene de neurofibromatose, que tinha sido perturbado pelo ponto de quebra da translocação. A análise subsequente mostrou que os pontos de quebra do cromossomo 17 não estavam

Pseudoligação de genes em uma translocação heterozigota

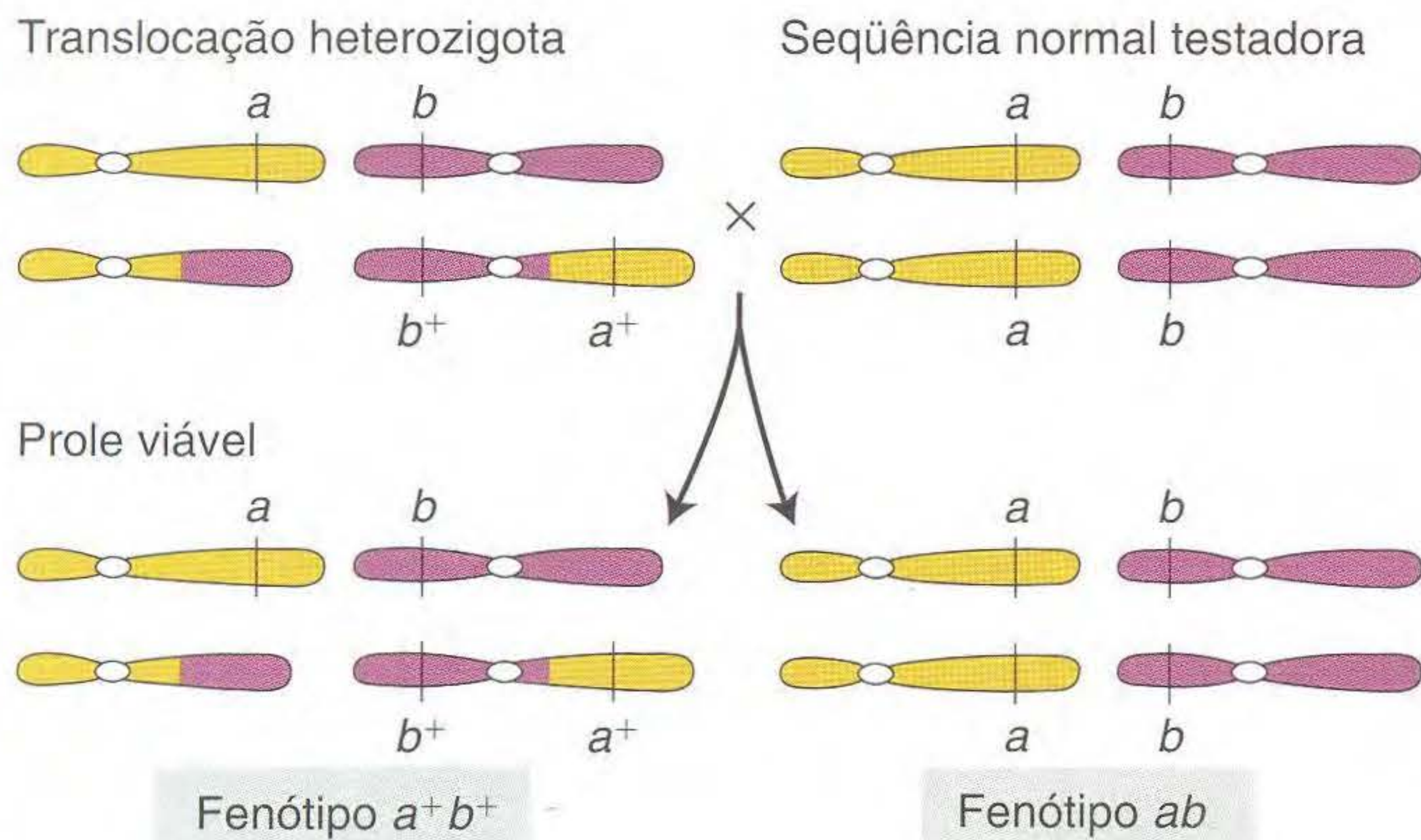


FIG. 16.32 Quando um fragmento translocado leva um gene marcador, esse marcador pode mostrar ligação com genes no outro cromossomo.



em posições idênticas; entretanto, como eles devem ter sido dentro do gene, a faixa de suas posições revelou o segmento do cromossomo que constituía o gene de neurofibromatose. O isolamento dos fragmentos de DNA dessa região acabou levando à recuperação do próprio gene.

Síntese de duplicações ou deleções específicas As translocações e inversões são rotineiramente usadas para deletar ou duplicar segmentos cromossômicos específicos. Lembre, por exemplo, que as inversões pericêntricas, bem como as translocações, geram produtos de meiose que contêm uma

duplicação e uma deleção (veja Figuras 16.29 e 16.30). Se o segmento duplicado ou deletado é muito pequeno, então os produtos meióticos de duplicação e deleção são equivalentes a duplicações ou deleções, respectivamente. As duplicações e deleções são úteis para uma variedade de aplicações experimentais, incluindo o mapeamento de genes e a variação da dosagem gênica para o estudo da regulação, como visto nas seções precedentes.

Outro enfoque para criar duplicações usa translocações *insercionais* unidirecionais, nas quais um segmento de um cromossomo é removido e inserido em outro. Em um hete-

Síndrome de Down na prole de uma translocação heterozigota

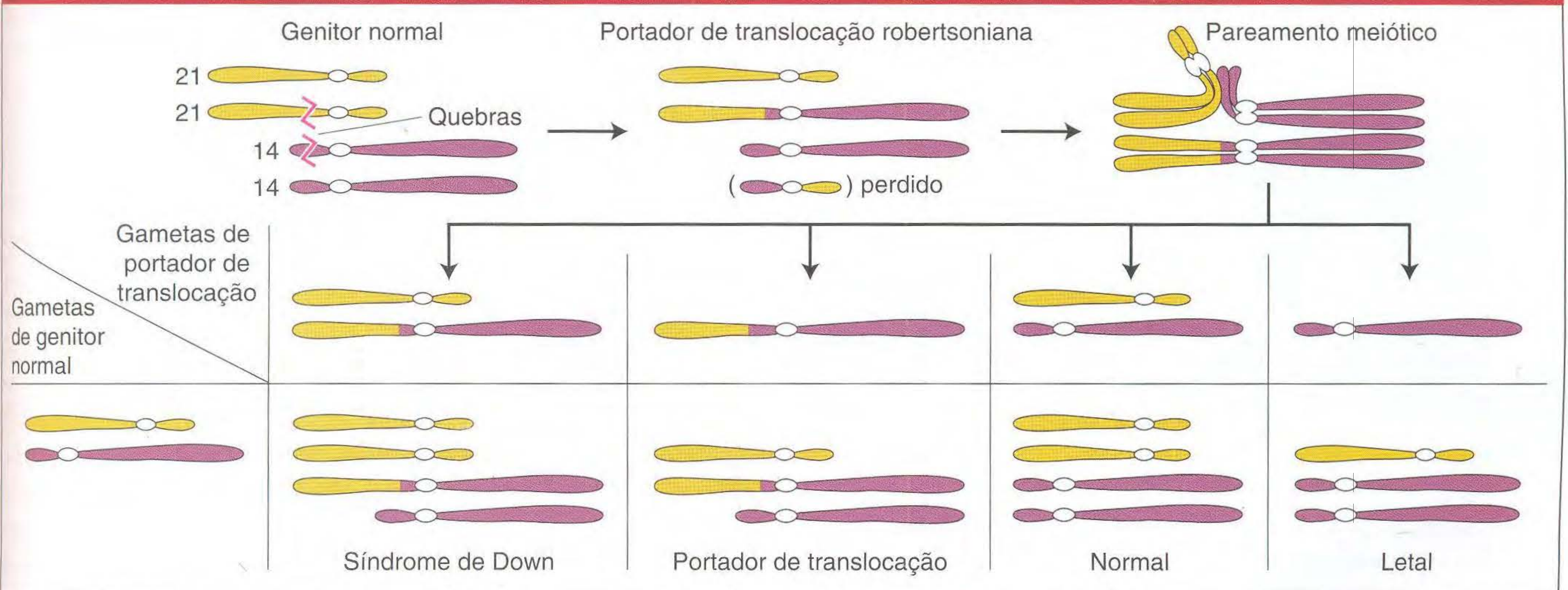


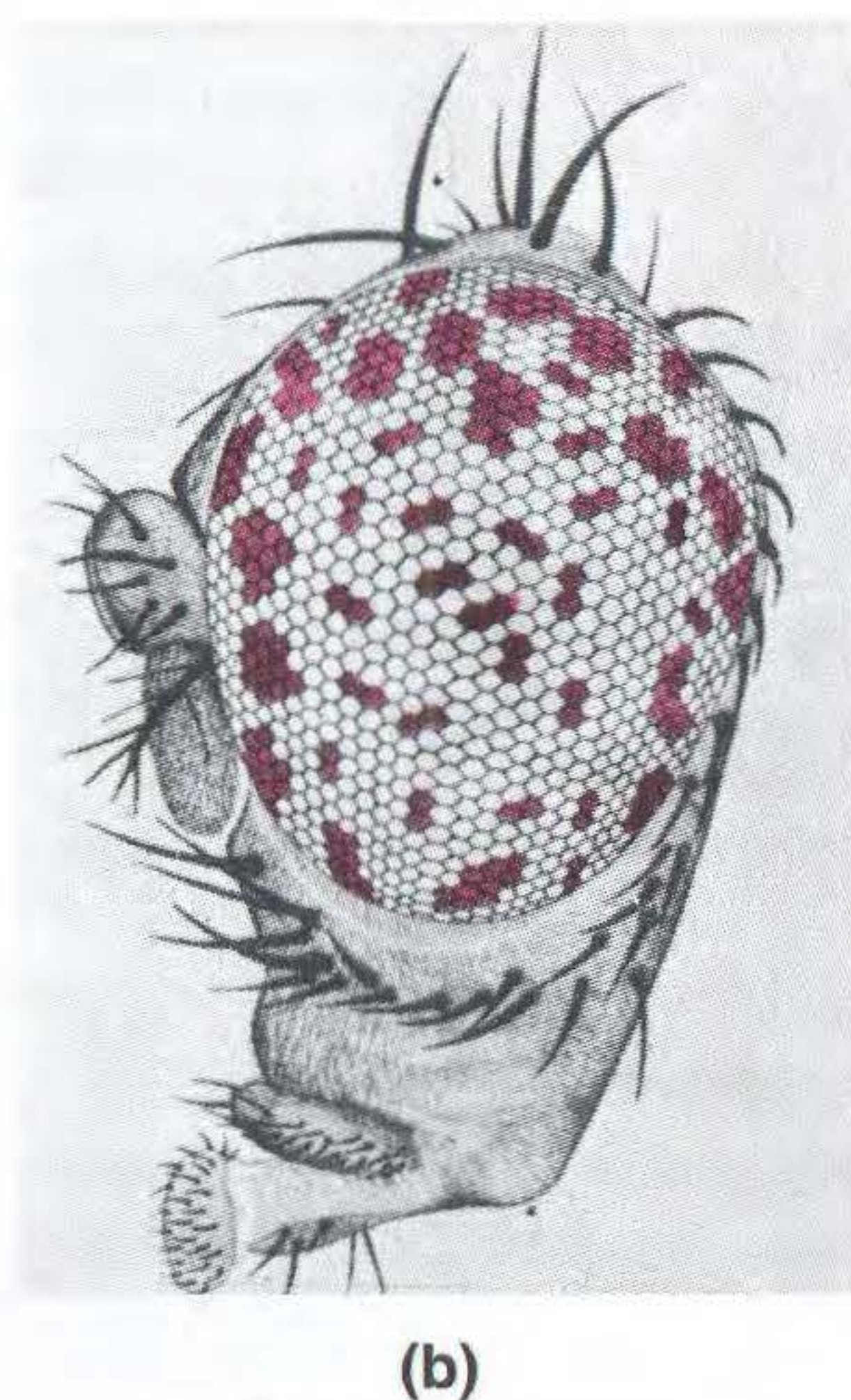
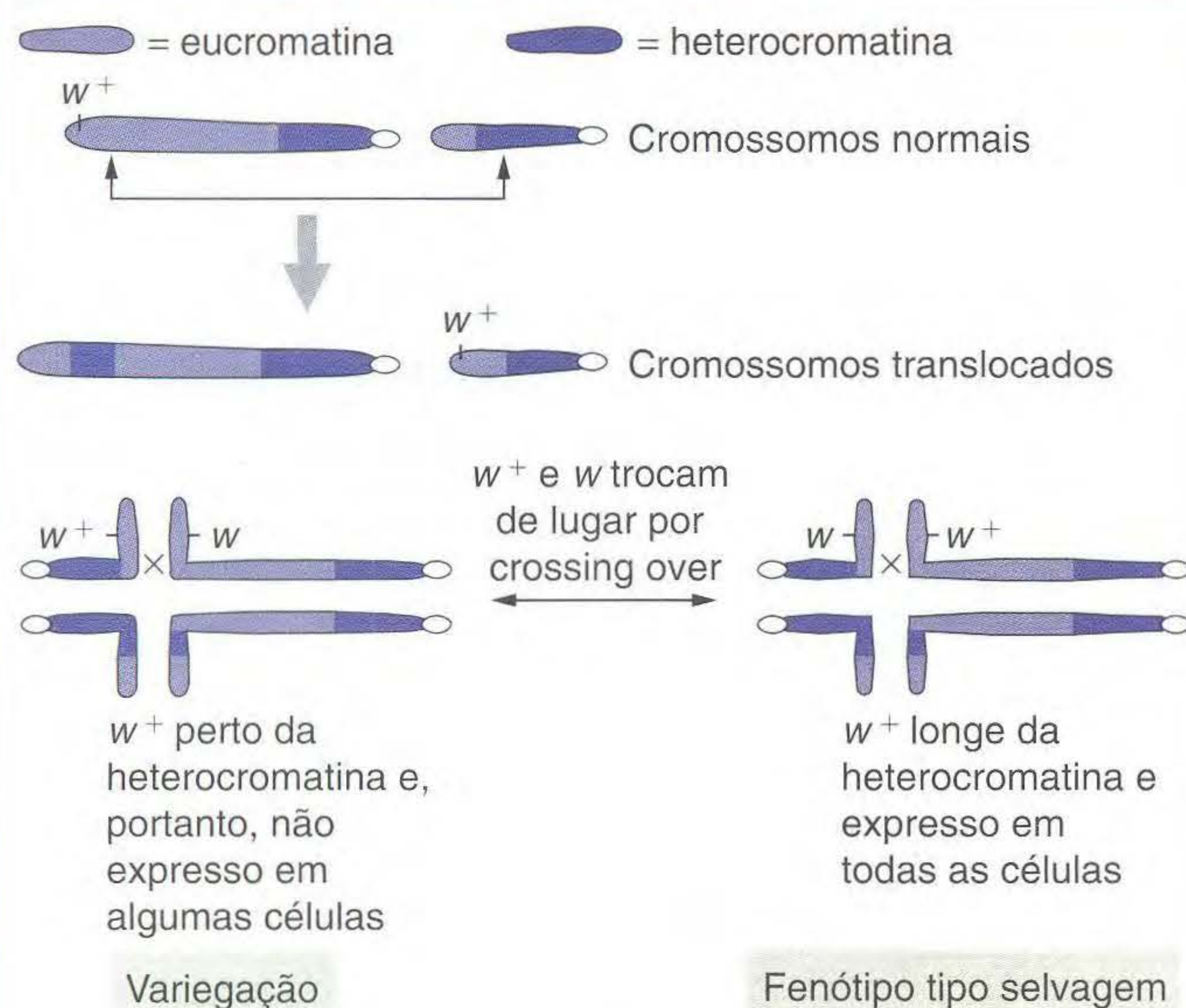
FIG. 16.33 Em uma minoria de casos, a origem da síndrome de Down é um genitor heterozigoto para uma translocação robertsoniana envolvendo o cromossomo 21. A segregação meiótica resulta em alguns gametas levando um cromossomo com um grande segmento adicional do cromossomo 21. Em combinação a um cromossomo 21 normal dado pelo gameta do sexo oposto, os sintomas da síndrome de Down são produzidos, muito embora não exista trissomia completa do 21.

rozigoto de translocação insercional, resulta uma duplicação se o cromossomo com a inserção segrega-se juntamente com a cópia normal.

Variegação de efeito de posição Como vimos no Capítulo 11, a ação gênica pode ser bloqueada pela proximidade a regiões cromossômicas densamente coradas, chamadas de *heterocromatina*. As translocações e inversões podem ser usadas para estudar esse efeito. Por exemplo, o locus para cor de olho branco em *Drosophila* está perto da ponta do cromossomo X. Considere uma translocação na qual a ponta de um cromossomo X levando w^+ é realocada para perto da região heterocromática, digamos, do cromossomo 4 (Figura 16.34a,

parte de cima). A **variegação de efeito de posição** é observada em moscas que são heterozigotas para tal translocação. O cromossomo X normal em tais heterozigotos portadores leva o alelo recessivo w . Espera-se que o fenótipo do olho seja vermelho, porque o alelo tipo selvagem é dominante em relação a w . Entretanto, em tais casos, o fenótipo observado é uma mistura variegada de facetas vermelhas e brancas (Figura 16.34b). Como podemos explicar as áreas brancas? O alelo w^+ nem sempre é expresso, pois os limites da heterocromatina são um tanto variáveis: em algumas células, ele engloba e inativa o gene w^+ , permitindo assim a expressão de w . Se as posições dos alelos w^+ e w são trocadas por um crossing, então a variegação do efeito de posição não é detectada (veja Figura 16.34a, parte de baixo).

A variegação pode resultar de uma proximidade incomum de um gene da heterocromatina em uma translocação heterozigota



Rearranjos e câncer



O câncer é uma doença de proliferação celular anormal. Como resultado de alguma agressão, uma célula do corpo se divide descontroladamente para formar uma população de células chamada câncer. Um aglomerado localizado de células proliferantes é chamado de um tumor, enquanto os cânceres de células móveis, tais como células sanguíneas, dispersam-se pelo corpo. O câncer é mais frequentemente causado por uma mutação na sequência codificante ou reguladora de um gene cuja função normal é regular a divisão celular. Tais genes são chamados de *proto-oncogenes*. Entretanto, os rearranjos cromossômicos, especialmente translocações, também podem interferir no funcionamento normal de tais proto-oncogenes.

Existem dois modos básicos pelos quais as translocações podem alterar a função dos proto-oncogenes. No primeiro mecanismo, a translocação realoca um proto-oncogene para perto de um novo elemento regulador. Um bom exemplo é dado pelo linfoma de Burkitt. O proto-oncogene nesse câncer codifica a proteína MYC, um fator de transcrição que ativa genes necessários para a proliferação celular. Normalmente, o gene *myc* é transcrito apenas quando uma célula precisa sofrer proliferação, mas, nas células cancerosas, o proto-oncogene MYC é realocado perto da região regulatória dos genes de imunoglobulina (Ig) (Figura 16.35a). Esses genes de imunoglobulina são constitutivamente transcritos; isto é, eles estão ligados o tempo todo. Conseqüentemente, o gene *myc* é transcrito o tempo todo, e os genes de proliferação celular são continuamente ativados.

O outro mecanismo pelo qual as translocações podem causar câncer é a formação de um gene híbrido. Um exemplo é dado pela doença leucemia mieloide crônica (CML), um câncer de leucócitos. Esse câncer pode resultar da formação de um gene híbrido entre os proto-oncogenes *BCR1* e *ABL* (Figura 16.35b). O proto-oncogene *abl* codifica uma cinase proteica em uma via de sinalização. A cinase proteica transmite um sinal iniciado por um fator de crescimento que leva à proliferação celular. A proteína de fusão Bcr1-Abl tem uma atividade permanente de cinase proteica. A proteína alterada propaga continuamente seu sinal de crescimento, independentemente da presença do sinal de iniciação.

FIG. 16.34 (a) A translocação de w^+ para uma posição perto da heterocromatina faz com que o funcionamento de w^+ falhe em algumas células, produzindo variegação de efeito de posição. (b) Um olho de *Drosophila* mostrando variegação de efeito de posição. [(b) De Randy Mottus.]

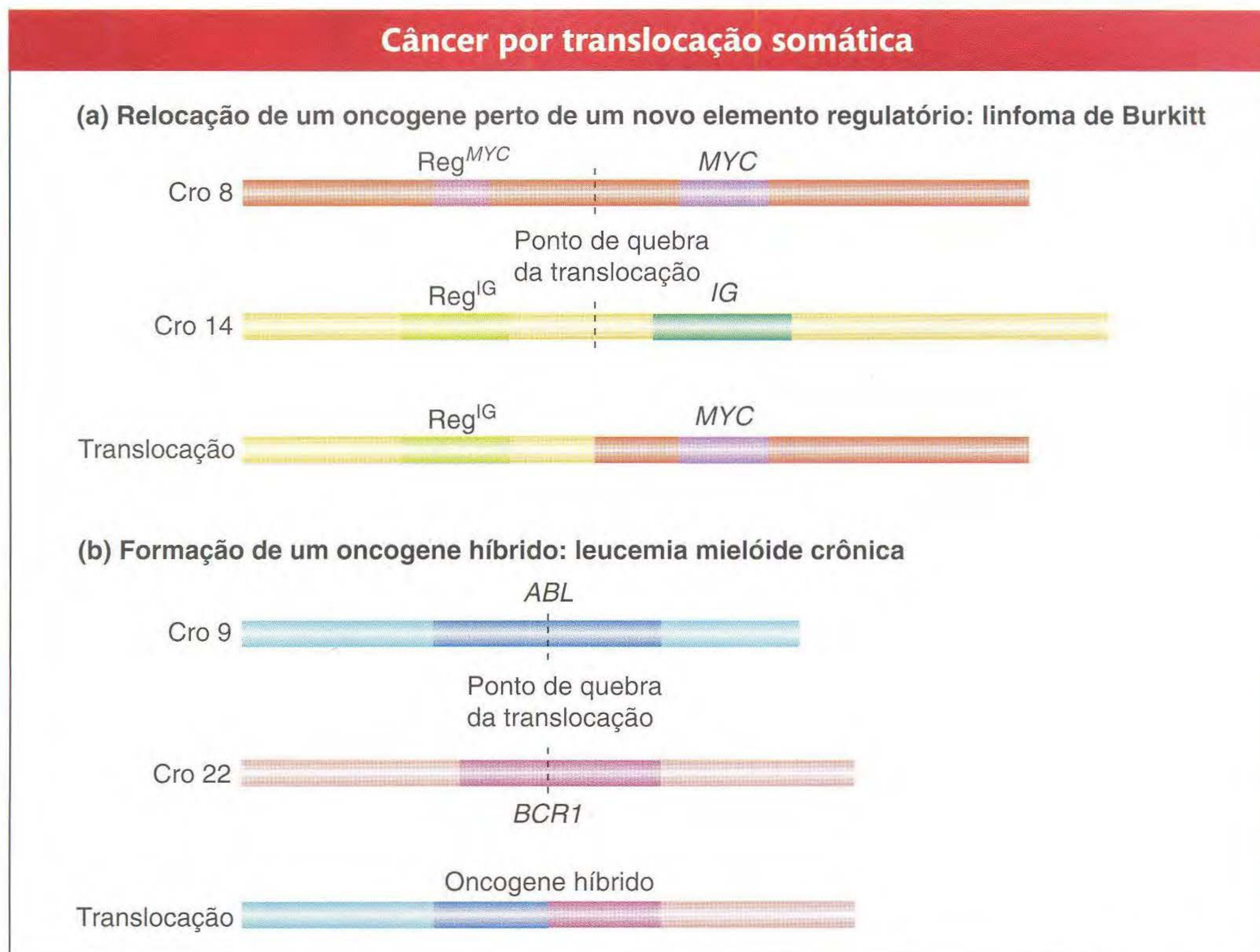


FIG. 16.35 Os dois principais modos pelos quais a translocação pode causar câncer em uma célula somática são ilustrados pelo linfoma de Burkitt (a) e a leucemia mielóide crônica (b). Os genes *MYC*, *BCR1* e *ABL* são proto-oncogenes.

Identificação das mutações cromossômicas pela genômica

Os microarranjos de DNA (veja Figura 13.16) possibilitaram detectar e quantificar as duplicações ou deleções de determinado segmento de DNA. A técnica é chamada de *hibridização genômica comparativa*. O DNA total do tipo selvagem e o de um mutante são marcados com dois corantes fluorescentes diferentes que emitem comprimentos de onda de luz distintos. Esses DNA marcados são adicionados a um microarranjo de cDNA, e ambos hibridizam-se a um arranjo. O arranjo é então escaneado por um detector ajustado para um comprimento de onda fluorescente, e é então novamente escaneado para outro comprimento de onda. A proporção de valores de cada cDNA é calculada. A proporção de mutantes para tipo selvagem substancialmente maior que 1 representa regiões que foram amplificadas. Uma proporção de 2 indica uma duplicação, e uma proporção de menos de 1 indica uma deleção. Alguns exemplos são mostrados na Figura 16.36.

16.3 Incidência Geral de Mutações Cromossômicas Humanas

As mutações cromossômicas são surpreendentemente frequentes na reprodução sexual humana, mostrando que os relevantes processos celulares são propensos a um alto nível de erro. A Figura 16.37 mostra a distribuição estimada de mutações cromossômicas entre as concepções humanas

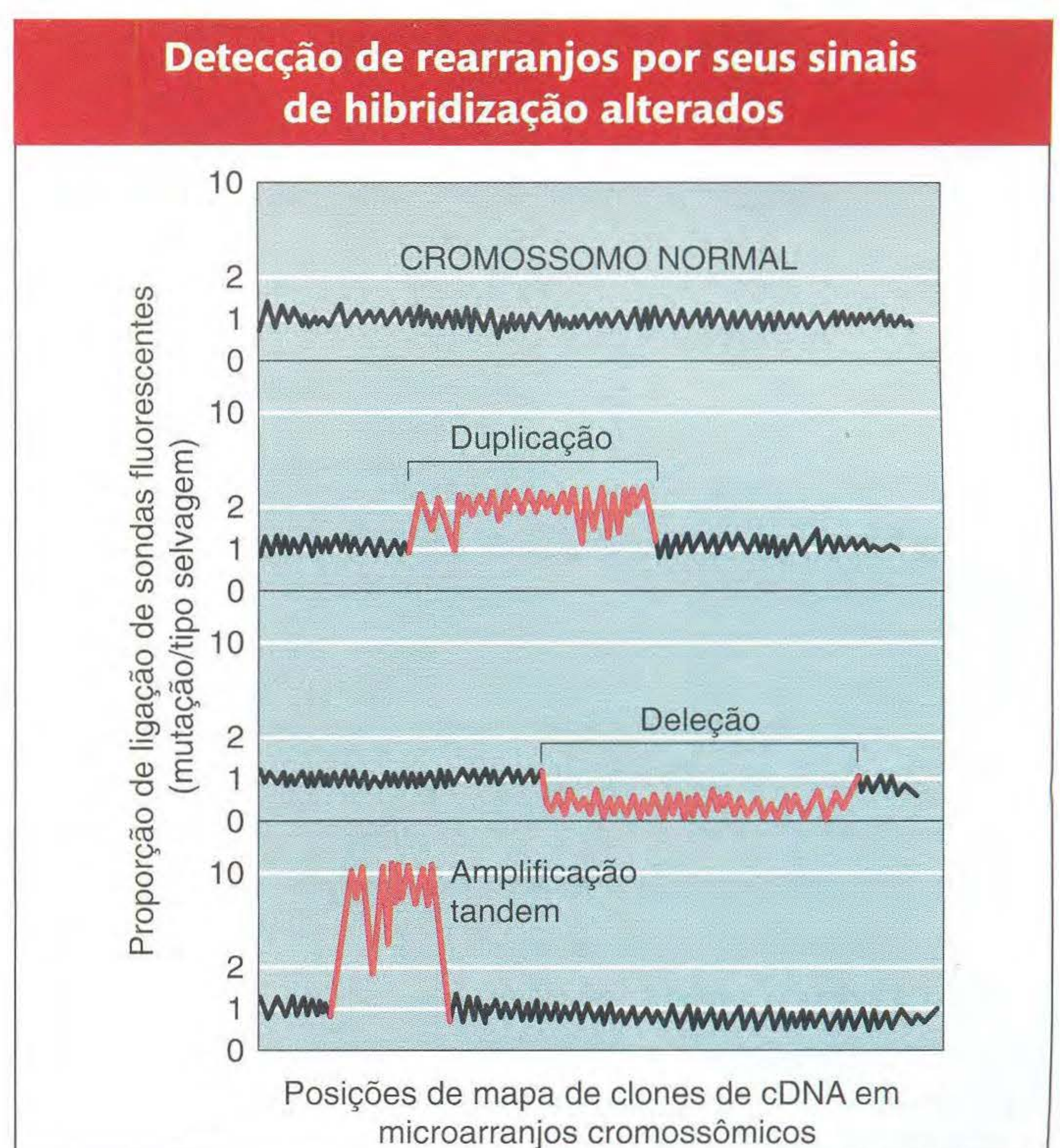


FIG. 16.36 Para detectar rearranjos cromossômicos, DNA genômicos mutantes e tipo selvagem são marcados com corantes que fluorescem em diferentes comprimentos de onda. Esses DNA marcados são adicionados a clones de cDNA dispostos em microarranjos ordenados cromossomicamente, e a proporção de fluorescência ligada em cada comprimento de onda é calculada para cada clone. São ilustrados os resultados esperados para um genoma normal e os três tipos de mutantes.

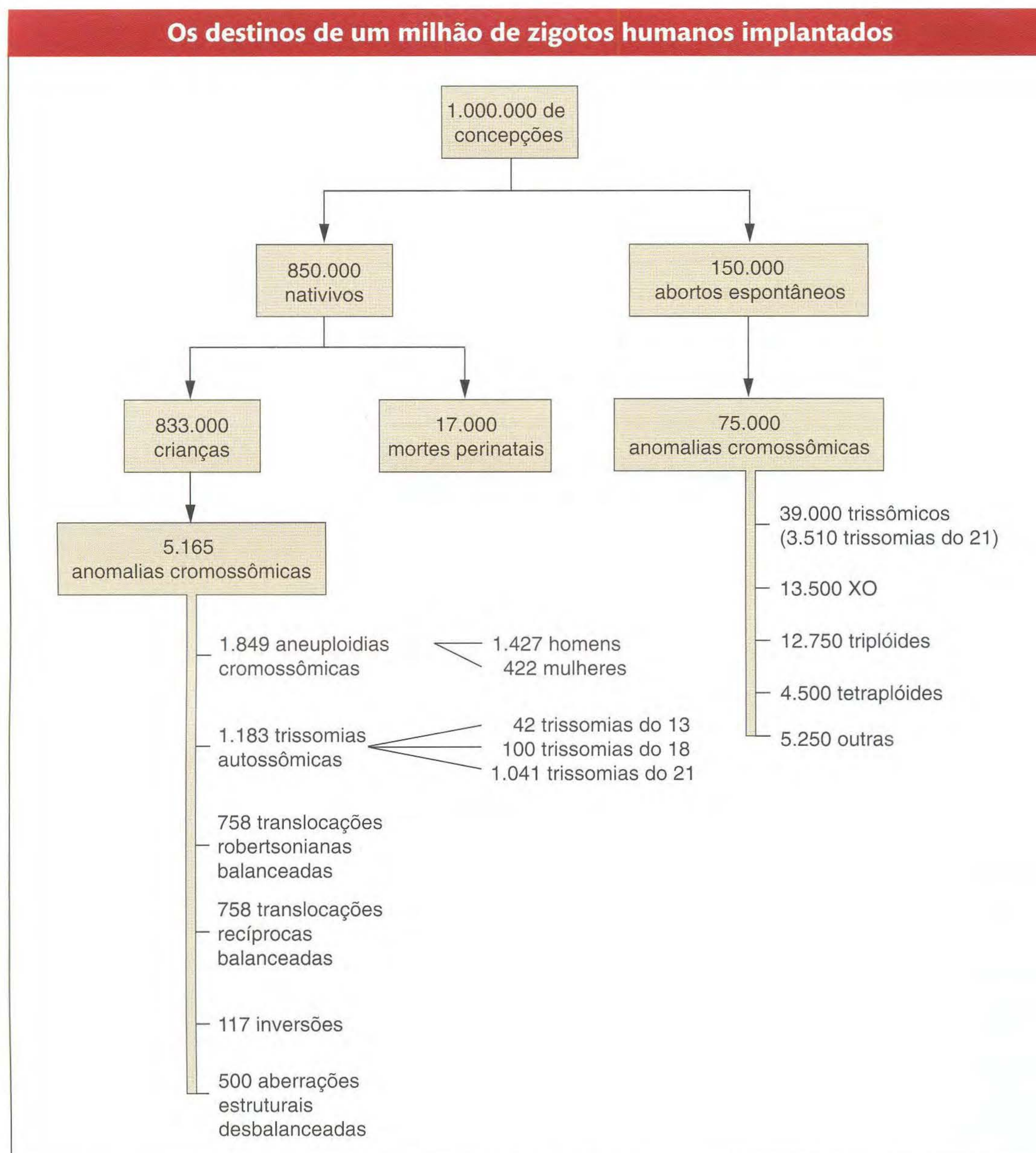


FIG. 16.37 A proporção de mutações cromossômicas é muito mais alta em abortos espontâneos. [De K. Sankaranarayanan, *Mutat. Res.* 61, 1979, 249-257.]

que se desenvolvem suficientemente para se implantar no útero. Dos 15% estimados de concepções que são abortadas espontaneamente (gestações que terminam naturalmente), metade apresenta anomalias cromossômicas. Alguns geneti-

cistas médicos acreditam que mesmo esse alto nível é uma subestimativa, pois muitos casos nunca são detectados. Entre os nativos, 0,6% têm anomalias cromossômicas, resultantes tanto de aneuploidias quanto de rearranjos cromossômicos.

► Resumo

A poliploidia é uma condição anormal na qual há um número maior do que o normal de conjuntos cromossômicos. Os poliplóides tais como os triplóides ($3n$) e os tetraplóides ($4n$) são comuns entre as plantas, e são representados mesmo entre animais. Os organismos com um número ímpar de conjuntos de cromossomos são estéreis porque nem todos os cromossomos têm um parceiro na meiose. Os cromossomos não-pareados ligam-se aleatoriamente aos pólos da célula na

meiose, levando a conjuntos desbalanceados de cromossomos nos gametas resultantes. Tais gametas desbalanceados não produzem prole viável. Nos poliplóides com número par de conjuntos, cada cromossomo tem um parceiro potencial de pareamento e, portanto, pode produzir gametas balanceados e uma prole. A poliploidia pode resultar em um organismo de maiores dimensões; essa descoberta permitiu avanços importantes na horticultura e na criação de cultivos.

Nas plantas, os alopoliplóides (poliplóides formados de combinações de conjuntos cromossômicos de espécies diferentes) podem ser feitos cruzando duas espécies relacionadas e, então, duplicando os cromossomos da prole com o uso de colchicina ou pela fusão de células somáticas. Essas técnicas têm aplicações potenciais nos cruzamentos de cultivos porque os alopoliplóides combinam as características de duas espécies parentais.

Quando acidentes celulares mudam partes de conjuntos cromossômicos, resultam os aneuplóides. A aneuploidia em si em geral resulta em um genótipo desbalanceado com um fenótipo anormal. Os exemplos de aneuplóides incluem monossômicos ($2n - 1$) e trissômicos ($2n + 1$). As síndromes de Down (trissomia do 21), de Klinefelter (XXY) e de Turner (XO) são exemplos bem documentados de condições aneuplóides em humanos. O nível espontâneo de aneuploidia em humanos é bem alto, e contribui com uma grande proporção de distúrbios geneticamente baseados em populações humanas. O fenótipo de um organismo aneuplóide depende muito do cromossomo particularmente afetado. Em alguns casos, tais como a trissomia do 21 humana, há uma constelação altamente característica de fenótipos associados.

A maioria dos casos de aneuploidia resulta de má segregação cromossômica acidental na meiose (não-disjunção). O erro é espontâneo e pode ocorrer em qualquer meiócito na primeira ou segunda divisão. Em humanos, um efeito de idade materna está associado à não-disjunção do cromossomo 21, resultando em uma maior incidência de síndrome de Down nos filhos de mães com mais idade.

A outra categoria geral de mutações cromossômicas compreende rearranjos estruturais, que incluem deleções, duplicações, inversões e translocações. Essas mudanças resultam ou de quebra e reunião incorreta, ou de crossing over entre elementos repetitivos (recombinação homóloga não-alélica). Os rearranjos cromossômicos são uma causa importante de problemas de saúde em populações humanas, e são úteis em construir linhagens especiais de organismos para genética experimental e aplicada. Em organismos com um conjunto cromossômico normal mais um conjunto rearranjado (rearranjos heterozigotos), existem estruturas incomumente pareadas na meiose que resultam de forte afinidade de pareamento de regiões cromossômicas homólogas. Por exemplo, as inversões heterozigotas apresentam alças, e translocações

recíprocas mostram estruturas em forma de cruz. A segregação dessas estruturas resulta em produtos meióticos anormais únicos para o rearranjo.

Uma deleção é a perda de uma seção do cromossomo, seja devido a quebra cromossômica seguida de uma perda do segmento intercalar, seja devido a segregação em translocações heterozigotas ou inversões. Se a região removida em uma deleção é essencial para a vida, a deleção homozigota é letal. As deleções heterozigotas podem ser letais devido ao desequilíbrio cromossômico ou porque revelam alelos deletérios recessivos, ou podem ser não-letais. Quando uma deleção em um homólogo permite a expressão fenotípica de alelos recessivos no outro, o desmascaramento dos alelos recessivos é chamado de pseudodominância.

As duplicações geralmente são produzidas por outros rearranjos ou por crossing over aberrante. Elas também desequilibram o material genético, produzindo um efeito fenotípico deletério ou a morte do organismo. Entretanto, as duplicações podem ser uma fonte de novo material para evolução porque a função pode ser mantida em uma cópia, deixando a outra cópia livre para desenvolver novas funções.

Uma inversão é um giro de 180° de uma parte de um cromossomo. No estado homozigoto, as inversões podem causar pouco problema para um organismo, a menos que a heterocromatina cause um efeito de posição ou que uma das quebras perturbe um gene. Por outro lado, as inversões heterozigotas apresentam alças de inversão na meiose, e o crossing over dentro das alças resultam em produtos inviáveis. Os produtos de crossing de inversões pericêntricas, que envolvem o centrômero, diferem das inversões paracêntricas, que não envolvem, mas ambas apresentam reduzida frequência de recombinantes na região afetada e, geralmente, resultam em reduzida fertilidade.

Uma translocação move um segmento cromossômico para outra posição no genoma. Um exemplo simples é uma translocação recíproca, na qual partes dos cromossomos não-homólogos trocam posições. No estado heterozigoto, as translocações produzem duplicações e deleções meióticas, que podem levar a zigotos desbalanceados. Podem ser produzidas novas ligações genéticas por translocações. A segregação aleatória de centrômeros em uma translocação heterozigota resulta em 50% de produtos meióticos desbalanceados e, assim, 50% de esterilidade (semi-esterilidade).

Termos-chave

alça de deleção	cromossomo acêntrico	duplicação segmentar
alça de inversão	cromossomo dicêntrico	efeito de dosagem gênica
alopoliplóide	cromossomo politênico	embrióide
aneuplóide	cromossomos homeólogos	euplóide
anfidiplóide	deleção	fragmento acêntrico
autopoliplóide	deleção intragênica	hexaplóide
balanceador	deleção multigênica	inversão
balanço gênico	dissômico	inversão heterozigota
bivalente	duplicação	inversão paracêntrica
carga genética	duplicação em tandem	inversão pericêntrica
compensação de dose	duplicação insercional	mapeamento de deleção

monoplóide
monossômico
mutação cromossômica
não-disjunção
nulissômicos
partenogênese
pentaplóide
poliplóide
ponte anafásica
ponte dicêntrica

pseudodominância
pseudoligação
rearranjo
rearranjo balanceado
rearranjo desbalanceado
recombinação homóloga não-alélica (NAHR)
segregação adjacente-1
segregação alternada
semi-esterilidade

síndrome de Down
síndrome de Klinefelter
síndrome de Turner
tetraplóide
translocação
triplóide
trissômico
trivalente
univalente
variação de efeito de posição

► Problemas Resolvidos

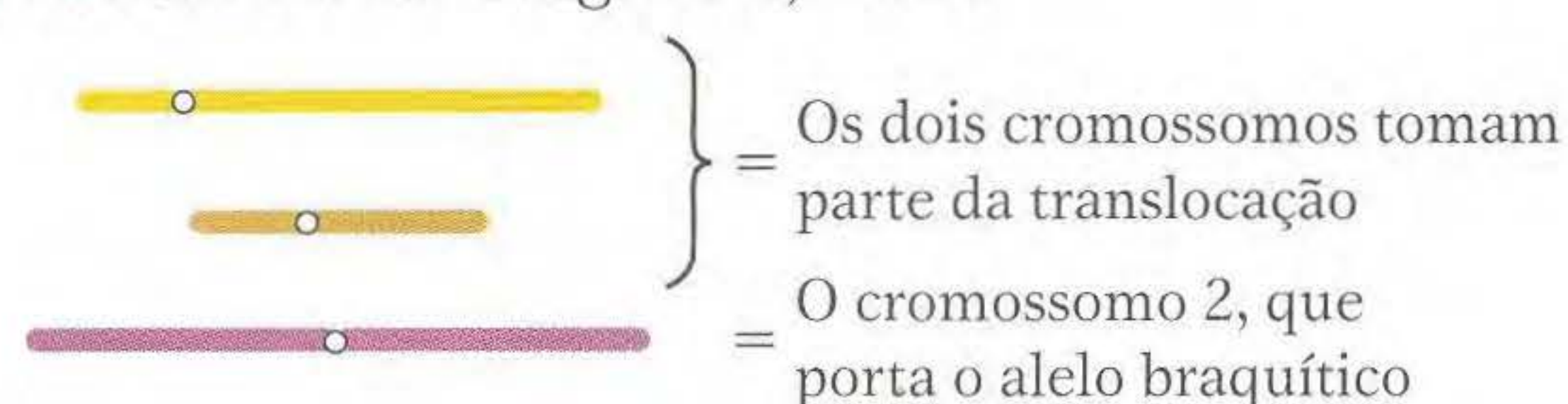
Problema resolvido 1. Uma planta de milho é heterozigota para uma translocação recíproca e, portanto, estéril. Essa planta é cruzada com uma linhagem cromossomicamente normal que é homozigota para o alelo recessivo braquítico (b), situado no cromossomo 2. Uma planta F_1 semi-estéril é então retrocruzada com a linhagem homozigota braquítica. A prole obtida tem os seguintes fenótipos:

Não-braquítico		Braquítico	
Semi-estéril	Fértil	Semi-estéril	Fértil
334	27	42	279

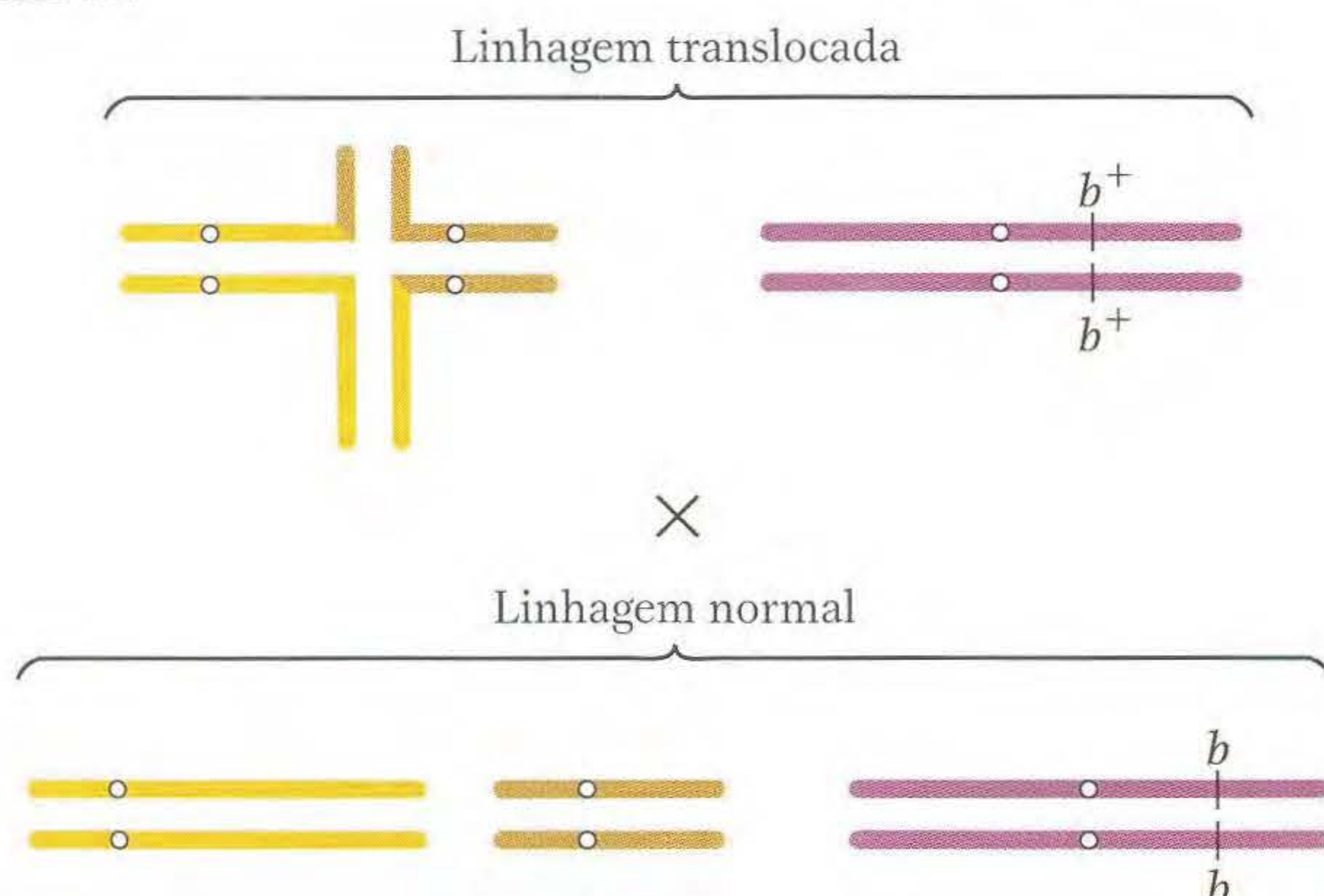
- Que proporção você espera resultar se o cromossomo que leva o alelo braquítico não toma parte da translocação?
- Você acha que o cromossomo 2 toma parte da translocação? Explique sua resposta, mostrando a conformação dos cromossomos relevantes da F_1 semi-estéril e o motivo dos números específicos obtidos.

SOLUÇÃO

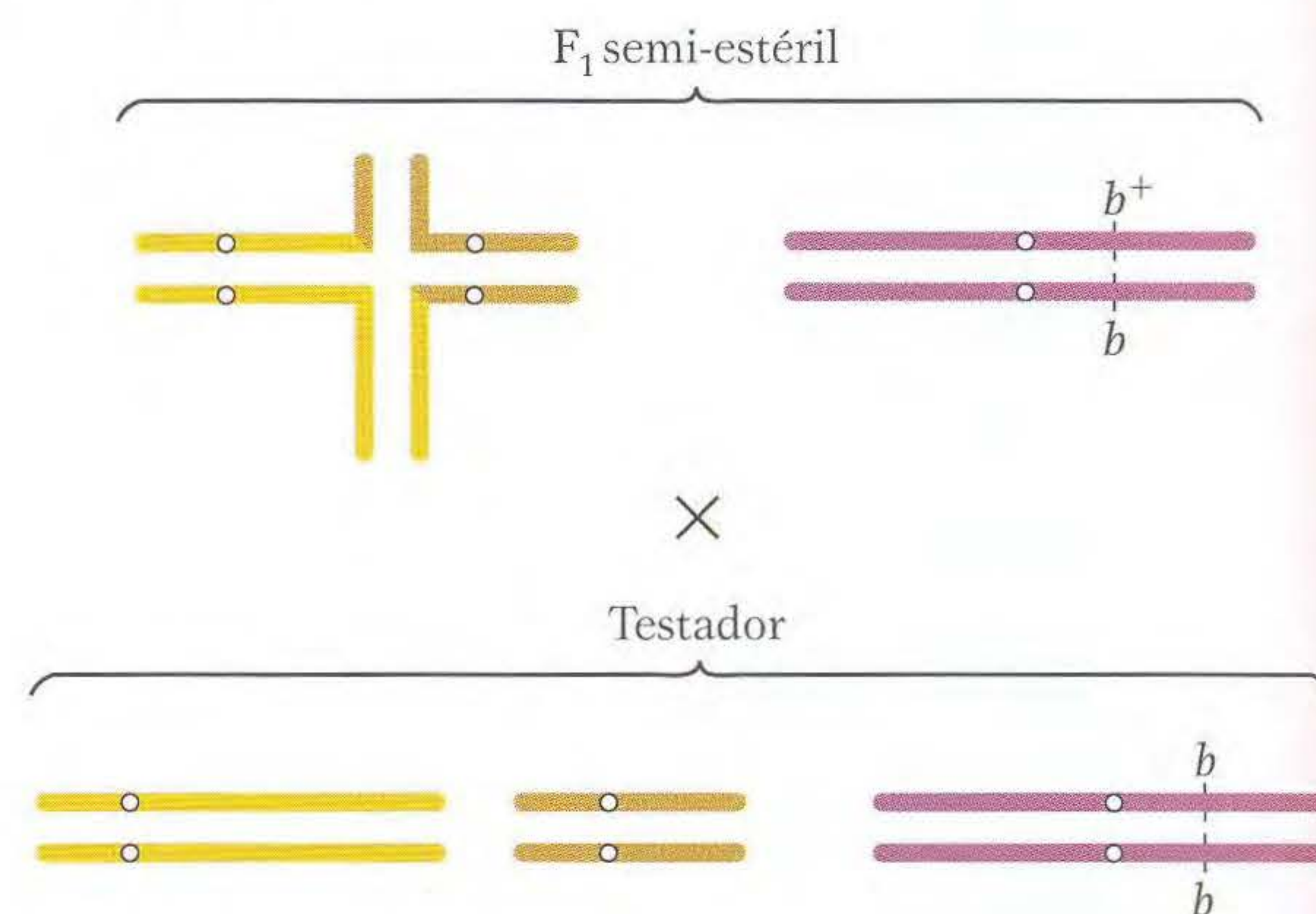
a. Devemos começar com o enfoque metódico e colocar os dados em forma de diagrama, onde



Para simplificar o diagrama, não mostramos os cromossomos divididos em cromátides (embora devam estar nesse estágio da meiose). Fizemos então um diagrama do primeiro cruzamento:



Toda a prole desse cruzamento será heterozigota para o cromossomo levando o alelo braquítico; mas, e quanto aos cromossomos que tomam parte na translocação? Neste capítulo, vimos que apenas os produtos de segregação alternada sobrevivem, e que metade desses sobreviventes serão cromossomicamente normais e metade portarão os cromossomos rearranjados. A combinação rearranjada regenerará um heterozigoto para a translocação quando ele se combina com o complemento cromossomicamente normal do genitor normal. Esses últimos tipos, os semi-estéreis de F_1 , são diagramados como parte do retrocruzamento com a linhagem parental braquítica:



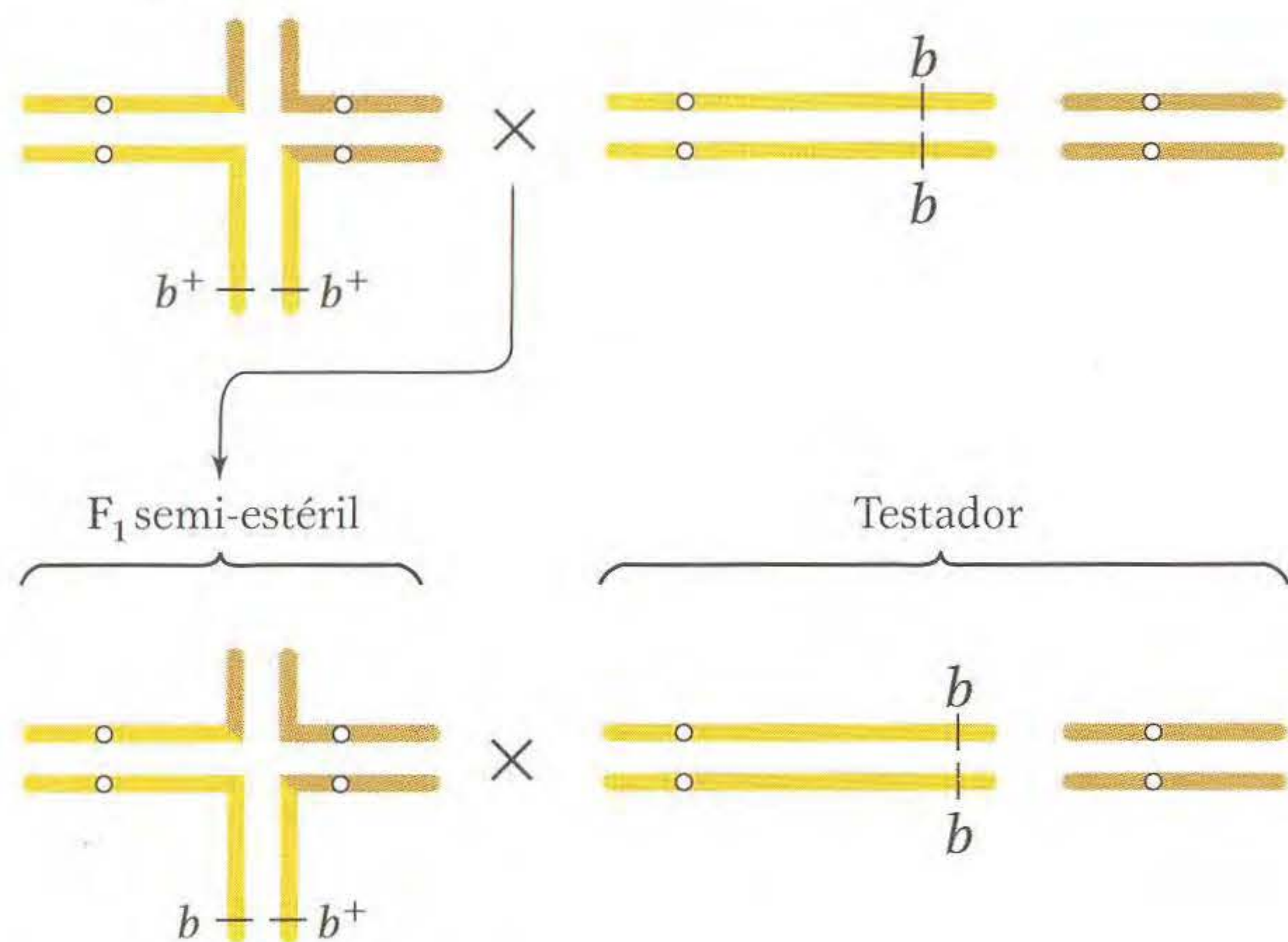
Calculando a proporção esperada de fenótipos desse cruzamento, podemos tratar o comportamento dos cromossomos translocados independentemente do comportamento do cromossomo 2. Assim, podemos prever que a prole será

$$\begin{array}{lcl}
 \frac{1}{2} \text{ heterozigotos de translocação (semi-estéreis)} & \begin{cases} \rightarrow \frac{1}{2} b^+/b \\ \rightarrow \frac{1}{2} b/b \end{cases} & \begin{cases} \rightarrow \frac{1}{4} \text{ semi-estéreis não-braquíticos} \\ \rightarrow \frac{1}{4} \text{ semi-estéreis braquíticos} \end{cases} \\
 \frac{1}{2} \text{ normais (férteis)} & \begin{cases} \rightarrow \frac{1}{2} b^+/b \\ \rightarrow \frac{1}{2} b/b \end{cases} & \begin{cases} \rightarrow \frac{1}{4} \text{ férteis não-braquíticos} \\ \rightarrow \frac{1}{4} \text{ férteis braquíticos} \end{cases}
 \end{array}$$

Essa proporção prevista de 1:1:1:1 é bem diferente da obtida no cruzamento real.

b. Como observamos um desvio da proporção esperada com base na independência do fenótipo braquítico e a

semi-esterilidade, o cromossomo 2 provavelmente toma parte da translocação. Suponhamos que o locus braquítico (b) está no cromossomo laranja. Mas onde? Para fins de diagrama não importa onde o pomos, mas importa sim, geneticamente, porque a posição do locus b afeta as proporções da prole. Se for suposto que o locus b está perto da ponta do pedaço que é translocado, podemos redesenhar o heredograma:

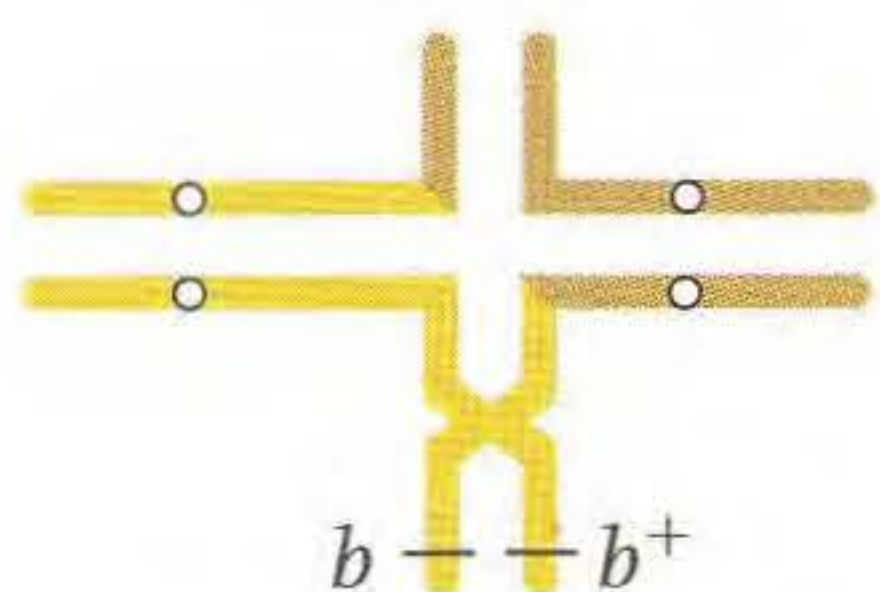


Se os cromossomos da F_1 semi-estéril segregarem como diagramado aqui, podemos então prever

$\frac{1}{2}$ fértil, braquítica

$\frac{1}{2}$ semi-estéril, não-braquítica

A maior parte da prole certamente é desse tipo, e, assim, devemos estar no caminho certo. Como são produzidos os dois tipos menos freqüentes? De algum modo, temos que ter o alelo b^+ no cromossomo laranja normal e o alelo b no cromossomo translocado. Esse posicionamento deve ser obtido por crossing over entre o ponto de quebra da translocação (o centro da estrutura em forma de cruz) e o locus braquítico.



Os cromossomos recombinantes produzem alguma prole que é fértil e não-braquítica e algumas que são semi-estéreis e braquíticas (essas duas classes juntas constituem 69 de um total de prole de 682, ou uma freqüência de cerca de 10%). Podemos ver que essa freqüência é de fato uma medida da distância de mapa (10 u.m.) do locus braquítico do ponto de quebra. (O mesmo resultado básico teria sido obtido se tivéssemos desenhado o locus braquítico na parte do cromossomo do outro lado do ponto de quebra.)

Problema resolvido 2. Temos linhagens de camundongo homozigotas para dois fenótipos comportamentais alternativos que sabemos que são determinados por dois alelos em um único locus: v faz com que o camundongo tenha um andar “valsador”, enquanto V determina um andar

normal. Após cruzar os valsadores com os normais, observamos que a maior parte da F_1 é normal, mas, inesperadamente, há uma fêmea valsadora. Cruzamos a F_1 valsadora com dois machos valsadores diferentes e notamos que ela só produz prole valsadora. Quando a cruzamos com machos normais, ela produz prole normal e não-valsadora. Cruzamos três de sua prole feminina normal com dois de seus irmãos, e esses camundongos produziram uma prole de 60, todos normais. Quando, entretanto, cruzamos uma dessas mesmas três fêmeas com um terceiro irmão, temos seis normais e dois valsadores em uma ninhada de oito. Pensando nos genitores da F_1 valsadora, podemos considerar algumas possíveis explicações para esses resultados:

a. Um alelo dominante pode ter mutado para um alelo recessivo no genitor normal.

b. Em um genitor, pode ter havido uma mutação dominante em um segundo gene para criar um alelo epistático que atua para evitar a expressão de V , levando a valsador.

c. Uma não-disjunção meiótica do cromossomo levando V em seu genitor normal pode ter dado origem a um aneuplóide viável.

d. Pode ter havido uma deleção viável envolvendo V no meiócito do genitor normal.

Quais dessas explicações são possíveis, e quais são eliminadas pela análise genética? Explique em detalhe.

SOLUÇÃO

O melhor modo de responder a pergunta é tomar as explicações uma de cada vez e ver se cada uma se ajusta aos resultados dados.

a. Mutação V para v

Essa hipótese requer que a fêmea excepcional valsadora seja homozigota v/v . Essa suposição é compatível com os resultados de cruzar ambas com machos valsadores, os quais, se ela é v/v , produziram toda a prole valsadora (v/v), e com machos normais, que produziram toda a prole normal (V/v). Entretanto, os cruzamentos entre irmãos dentro dessa prole normal deveria então produzir uma proporção de 3:1 normais para valsadores. Como alguns dos cruzamentos entre irmãos de fato não produzem valsadores, essa hipótese não explica os dados.

b. Mutação epistática s para S

Aqui os genitores seriam $V/V \cdot s/s$ e $v/v \cdot s/s$, e uma mutação germinativa em um deles daria a F_1 valsadora com o genótipo $V/v \cdot S/s$. Quando a cruzamos com um macho valsador, que teria o genótipo $v/v \cdot s/s$, esperaríamos alguma prole $V/v \cdot S/s$, que seria fenotipicamente normal. Entretanto, não vimos prole normal desse cruzamento, e, assim, a hipótese já foi desprezada. A ligação poderia salvar temporariamente a hipótese se for suposto que a mutação foi no genitor normal, dando um gameta VS . Então a F_1 valsadora seria VS/vs , e, se a ligação fosse bem forte, seriam produzidos poucos gametas Vs ou nenhum, tipos que são necessários para combinar os gametas vs do macho para dar Vs/vs normais. Entretanto, se a hipótese de ligação fosse verdadeira, o cruzamento

com machos normais seria $Vs/vs \times Vs/Vs$, e isso daria uma alta percentagem de prole VS/Vs , que seriam valsadores, nenhum dos quais foi visto.

c. Não-disjunção no genitor normal

Essa explicação daria um gameta nulissômico que se combinaria com v para dar a F_1 valsadora com o genótipo v . As reproduções subseqüentes seriam

- $v \times v/v$, que daria prole v/v e v , todos valsadores. Isso se ajusta.
- $v \times V/V$, que daria prole V/v e V , todos normais. Isso também se ajusta.
- Primeiros entrecruzamentos de prole normal: $V \times V$. Esses entrecruzamentos dariam V e V/V , que são normais. Isso se ajusta.
- Segundos entrecruzamentos de prole normal: $V \times V/v$. Esses entrecruzamentos dão 25% cada de V/V , V/v , V (todos normais) e v (valsadores). Isso também se ajusta.

Essa hipótese é, portanto, consistente com os dados.

d. Deleção de V em um genitor normal

Chamemos a deleção de D . A F_1 valsadora seria D/v , e as reproduções subseqüentes seriam:

- $D/v \times v/v$, que dá v/v e D/v , que são valsadores. Isso se ajusta.
- $D/v \times V/V$, que dá V/v e D/V , que são normais. Isso se ajusta.
- Primeiros entrecruzamentos de prole normal: $D/V \times D/V$, que dá D/V e V/V , todos normais. Isso se ajusta.
- Segundos entrecruzamentos de prole normal: $D/V \times V/v$, que dá 25% cada de V/V , V/v , D/V (todos normais) e D/v (valsadores). Isto também se ajusta.

Novamente, a hipótese se ajusta aos dados fornecidos; logo, ficamos com duas hipóteses que são compatíveis com os resultados, e são necessários outros experimentos para distingui-los. Um modo de fazer isso seria examinar os cromossomos da fêmea excepcional ao microscópio. A aneuploidia deve ser fácil de distinguir da deleção.

► Problemas

PROBLEMAS BÁSICOS

1. Mantendo o estilo do Quadro 16.1, como você chamaria organismos que são MM N OO; MM NN OO; MMM NN PP?
2. Uma planta grande surgiu em uma população natural. Qualitativamente, ela parece igual, porém é muito maior. É mais provável ser um aloploplóide ou um autoploplóide? Como você testaria se foi poliploidia e não apenas o crescimento em solo rico?
3. Um trissômico é um aneuplóide ou um poliplóide?
4. Em um tetraplóide $B/B/b/b$, quantos possíveis pareamentos quadrivalentes existem? Desenhe-os (veja Figura 16.5).
5. Alguém lhe diz que a couve-flor é um anfidiplóide. Você concorda? Explique.
6. Por que *Raphanobrassica* é fértil, embora seu genitor não seja?
7. Na designação dos genomas de trigo, quantos cromossomos estão representados pela letra B?
8. Como você "recria" um trigo do pão hexaplóide a partir de *Triticum tauschii* e Emmer?
9. Como você faz uma plântula monoplóide começando com uma planta diplóide?
10. É obtido um produto dissômico da meiose. Qual sua provável origem? Que outros genótipos você espera entre os produtos dessa meiose sob sua hipótese?
11. Pode um trissômico $A/A/a$ produzir um gameta com o genótipo a ?
12. Quais, se algum, dos seguintes aneuplóides de cromossomos sexuais em humanos são férteis: XXX, XXY, XYY, XO?
13. Por que as mães com mais idade são rotineiramente examinadas com amniocentese ou CVS?
14. Em uma inversão, a ponta 5'DNA é unida alguma vez à ponta 5'? Explique.
15. Se você observa uma ponta dicêntrica na meiose, que rearranjo você pode prever que tenha ocorrido?
16. Por que os fragmentos acêntricos são perdidos?
17. Diagrame uma translocação surgindo de DNA repetitivo. Repita para uma deleção.
18. De um grande estoque de rearranjos disponíveis de *Neurospora* de um centro de estoques de genética de fungos, que tipo você escolheria para sintetizar uma linhagem que tinha uma duplicação do braço direito do cromossomo 3 e uma deleção da ponta do cromossomo 4?
19. Você observa uma alça de pareamento muito grande na meiose. É mais provável ser de uma inversão heterozigota ou deleção heterozigota? Explique.
20. Um novo alelo mutante recessivo mostra pseudodominância com qualquer das deleções que se encontram no cromossomo 2 de *Drosophila*. Qual pode ser a explicação?
21. Compare e contraste as origens das síndromes de Turner, de Williams, do *cri du chat* e de Down. (Por que elas são chamadas de *síndromes*?)
22. Liste as características diagnósticas (genéticas e citológicas) que são usadas para identificar estas alterações cromossômicas.
 - a. Deleção
 - b. Duplicações
 - c. Inversões
 - d. Translocações recíprocas
23. A seqüência normal de nove genes de determinado cromossomo de *Drosophila* é 123 · 456789, onde o ponto representa o centrômero. Algumas moscas-das-frutas foram encontradas tendo cromossomos aberrantes com as seguintes estruturas:

- a. 123 · 476589
- b. 123 · 46789
- c. 1654 · 32789
- d. 123 · 4566789

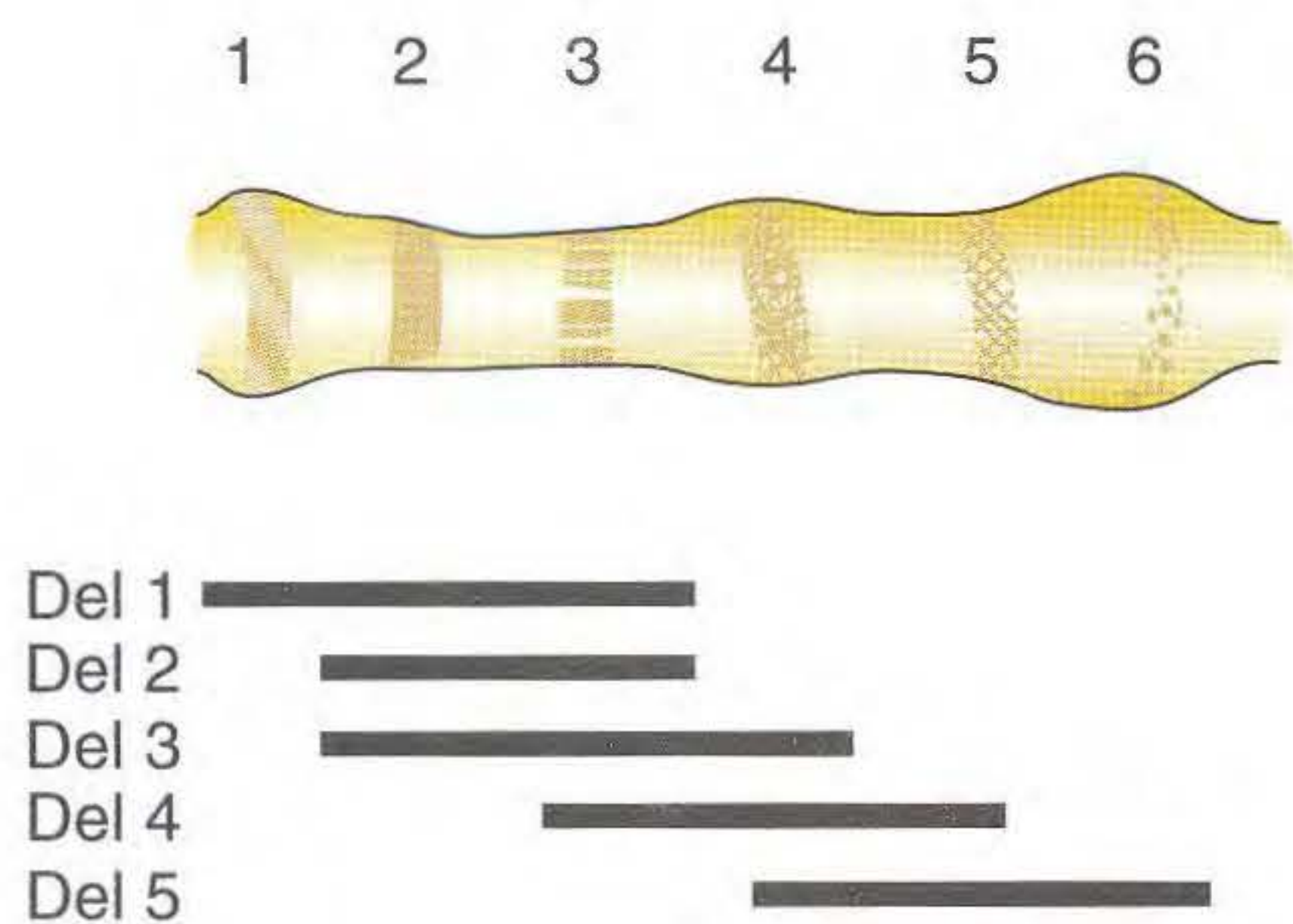
Nomeie cada tipo de rearranjo cromossômico, e desenhe diagramas para mostrar como cada um faria sinapse com o cromossomo normal.

24. Os dois loci *P* e *Bz* distam normalmente 36 u.m. no mesmo braço de um cromossomo de determinada planta. Certa inversão paracêntrica envolve mais ou menos um quarto dessa região, mas não inclui nenhum dos loci. Que frequência de recombinação aproximada entre *P* e *Bz* você poderia prever em plantas que são

- a. heterozigotas para a inversão paracêntrica?
- b. homozigotas para a inversão paracêntrica?

25. Como dito no Problema Resolvido 2, alguns camundongos chamados de *valsadores* têm uma mutação recessiva que faz com que eles executem passos bizarros. W.H. Gates cruzou valsadores com homozigotos normais e encontrou, entre várias centenas de prole normal, uma única fêmea valsadora. Quando cruzadas com um macho valsador, ela produziu uma prole toda valsadora. Quando cruzada com um macho normal homozigoto, produziu uma prole toda normal. Alguns machos e fêmeas dessa prole normal foram entecruzados, e não houve prole valsadora. T.S. Painter examinou os cromossomos dos camundongos valsadores que eram derivados dos cruzamentos de Gates e que apresentavam um comportamento reprodutivo similar ao da fêmea original valsadora. Ele observou que esses camundongos tinham 40 cromossomos, como nos camundongos normais ou nos valsadores incomuns. Entretanto, nos valsadores incomuns, um membro de um par de cromossomos era anormalmente pequeno. Interprete essas observações o mais completamente possível, tanto genética quanto citologicamente. (O problema 25 é de A.M. Srb, R.D. Owen, and R.S. Edgar, *General Genetics*, 2nd ed. W.H. Freeman and Company, 1965.)

26. Seis bandas em um cromossomo de glândula salivar de *Drosophila* são mostrados na ilustração seguinte, juntamente com cinco deleções (Del 1 a Del 5):



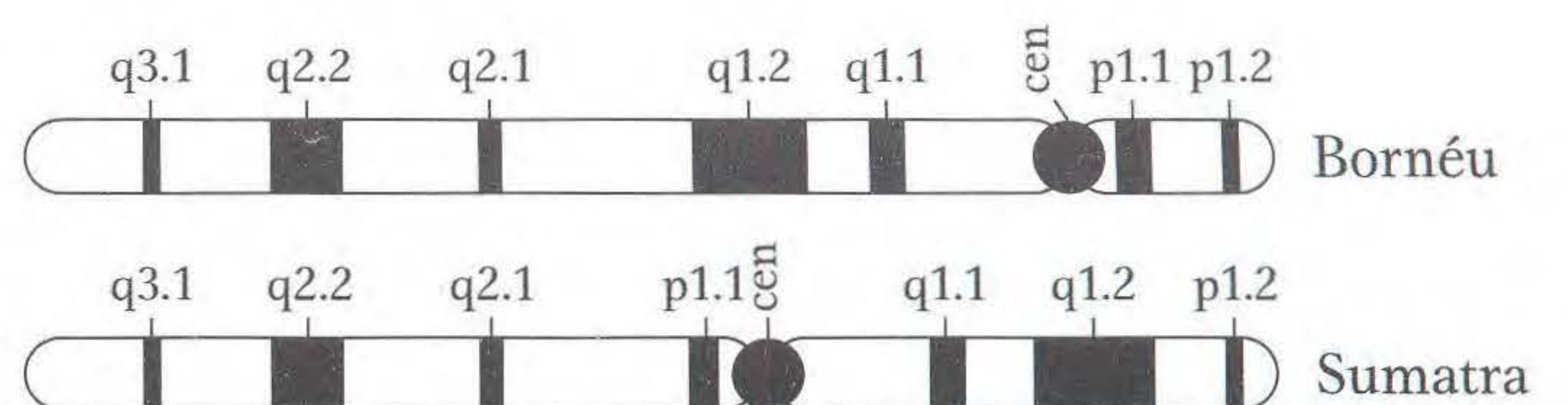
Os alelos recessivos *a*, *b*, *c*, *d*, *e* e *f* são conhecidos como estando na região, mas sua ordem é desconhecida. Quando as deleções são combinadas a cada alelo, são obtidos os seguintes resultados:

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>
Del 1	-	-	-	+	+	+
Del 2	-	+	-	+	+	+
Del 3	-	+	-	+	-	+
Del 4	+	+	-	-	-	+
Del 5	+	+	+	-	-	-

Nesse quadro, um sinal menos significa que a deleção não tem o alelo tipo selvagem correspondente (a deleção revela o recessivo), e um sinal mais significa que o alelo tipo selvagem correspondente ainda está presente. Use esses dados para deduzir que banda salivar contém cada gene. (O problema 26 é de D.L. Hartl, D. Friefelder, and L.A. Snyder, *Basic Genetics*. Jones and Bartlett, 1988.)

27. Foi encontrada uma mosca-das-frutas heterozigota para uma inversão paracêntrica. Entretanto, obter moscas que eram homozigotas para a inversão foi impossível mesmo após muitas tentativas. Qual a explicação mais provável para essa incapacidade de produzir uma inversão homozigota?

28. Os orangotangos são uma espécie em risco em seu ambiente natural (as ilhas de Bornéu e Sumatra), e, assim, foram estabelecidos programas de cruzamento em cativeiro usando orangotangos mantidos em zoos pelo mundo. Um componente desse programa é a pesquisa da citogenética dos orangotangos. Essa pesquisa mostrou que todos os orangotangos de Bornéu possuem uma forma de cromossomo 2, como mostrado no diagrama acompanhante, e todos os orangotangos de Sumatra têm outra forma. Antes que essa diferença citogenética fosse conhecida, foram feitos alguns cruzamentos entre animais de ilhas diferentes, e 14 de proles híbridas hoje são criados em cativeiro.



a. Que termo ou termos descrevem as diferenças entre esses cromossomos?

b. Desenhe os cromossomos 2, pareados na primeira prófase meiótica, de tal orangotango híbrido. Certifique-se de mostrar todos os marcros indicados no diagrama acompanhante e marque todas as partes em seu desenho.

c. Em 30% das meioses, haverá um crossing em algum lugar na região entre as bandas p1.1 e q1.2. Desenhe os cromossomos 2 dos gametas que resultarão de uma meiose na qual um único crossing ocorreu dentro da banda q1.1.

d. Que fração de gametas produzidos por um orangotango híbrido dará origem a uma prole viável, se esses cromossomos forem os únicos que diferem entre os genitores? (O problema 28 é de Rosemary Redfield.)

29. Em milho, os genes para o tamanho da borla (alelos T e t) e resistência à ferrugem (alelos R e r) são conhecidos estando em cromossomos separados. No curso dos cruzamentos rotineiros, um agricultor notou que uma planta $T/t ; R/r$ deu resultados incomuns em um cruzamento-teste com o pólen de um genitor duplo recessivo $t/t ; r/r$. Os resultados foram

Prole:	$T/t ; R/r$	98
	$t/t ; r/r$	104
	$T/t ; r/r$	3
	$t/t ; R/r$	5

Espigas: Apenas cerca de metade das sementes que o usual

- Quais as características-chave dos dados que são diferentes dos resultados esperados?
- Formule uma hipótese concisa que explique os resultados.
- Mostre os genótipos dos genitores e das proles.
- Desenhe um diagrama mostrando o arranjo dos alelos nos cromossomos.
- Explique a origem das duas classes de prole tendo três e cinco membros.



Solucionando o Problema 29

- O que significa um “gene para o tamanho da borla” e um “gene para resistência a ferrugem”?
- Tem importância se o significado preciso dos símbolos T , t , R e r não for dado? Por que sim ou por que não?
- Como os termos gene e alelo, usados aqui, se relacionam aos conceitos de locus e par de genes?
- Que evidência experimental anterior daria ao geneticista de milho a idéia de que os dois genes estão em cromossomos separados?
- O que você imagina que sejam “cruzamentos de rotina” para plantadores de milho?
- Que termo é usado para descrever genótipos do tipo $T/t ; R/r$?
- O que é um “genitor de pólen”?
- O que são cruzamentos-teste, e por que os geneticistas os acham úteis?
- Que tipos de prole e frequências os cultivadores esperam do cruzamento-teste?
- Descreva como a prole observada difere das expectativas.
- O que lhe diz a aproximada igualdade das primeiras duas classes de prole?
- O que lhe diz a aproximada igualdade das duas segundas classes de prole?
- Quais eram os gametas da prole incomum e quais as suas proporções?
- Que gametas estavam em maioria?
- Que gametas estavam em minoria?
- Quais os tipos de prole que pareciam ser recombinantes?
- Que combinações alélicas parecem estar ligadas de algum modo?
- Como pode haver ligação de genes supostamente em cromossomos separados?

- O que lhe dizem essas classes majoritárias e minoritárias sobre os genótipos dos genitores das plantas incomuns?
- O que é uma espiga de milho?
- Qual o aspecto de uma espiga de milho normal? (Desenhe uma e marque-a.)
- Quais os aspectos dessas espigas nesse cruzamento? (Desenhe uma.)
- O que é exatamente um grão?
- Que efeito pode levar à ausência de metade dos grãos?
- Metade dos grãos morreram? Caso sim, foi o genitor feminino ou masculino o motivo das mortes?

Agora tente resolver o problema.

30. Corpo amarelo em *Drosophila* é causado por um alelo mutante y de um gene situado na ponta do cromossomo X (o alelo tipo selvagem causa um corpo cinza). Em um experimento de radiação, um macho tipo selvagem foi irradiado com raios X e, então, cruzado com uma fêmea de corpo amarelo. A maioria da prole masculina era amarela, como esperado, mas a pesquisa de milhares de moscas revelou dois machos de corpo cinza (fenotipicamente tipo selvagem). Esses machos tipo selvagem foram cruzados com fêmeas de corpo amarelo, com os seguintes resultados:

	Prole
macho cinza 1 \times fêmea amarela	fêmeas todas amarelas machos todos cinza
macho cinza 2 \times fêmea amarela	$\frac{1}{2}$ fêmeas amarelas $\frac{1}{2}$ fêmeas cinza $\frac{1}{2}$ machos amarelos $\frac{1}{2}$ machos cinza

- Explique a origem e o comportamento do crossing do macho 1 cinza.
- Explique a origem e o comportamento do crossing do macho 2 cinza.



31. Em milho, o alelo Pr representa caules verdes, pr representa caules púrpura. Uma planta de milho com o genótipo pr/pr que tem cromossomos padrão é cruzada com uma planta Pr/Pr que é homozigota para uma translocação recíproca entre os cromossomos 2 e 5. A F_1 é semi-estéril e fenotipicamente Pr . Um retrocruzamento com o genitor que tem cromossomos padrão deu 764 semi-estéreis Pr , 145 semi-estéreis pr , 186 normais Pr e 727 normais pr . Qual a distância de mapa entre o locus Pr e o ponto de translocação?
32. Distinga as síndromes de Klinefelter, de Down e de Turner. Quais síndromes são encontradas em ambos os sexos?
33. Mostre como você pode fazer um alotetraplóide entre duas espécies de plantas diplóides correlatas, ambas as quais são $2n = 28$.

34. Em *Drosophila*, os trissômicos e monossômicos para o pequeno cromossomo 4 são viáveis, mas os nulissômicos e tetrassômicos não. O locus *b* está nesse cromossomo. Deduza as proporções fenotípicas na prole dos seguintes cruzamentos de trissômicos.
- $b^+/b/b \times b/b$
 - $b^+/b^+/b \times b/b$
 - $b^+/b^+/b \times b^+/b$
35. Uma mulher com síndrome de Turner é daltônica (um fenótipo recessivo ligado ao X). Tanto sua mãe quanto seu pai têm visão normal.
- Explique a origem simultânea da síndrome de Turner e o daltonismo pelo comportamento anormal de cromossomos na meiose.
 - Sua explicação pode distinguir se o comportamento cromossômico anormal ocorreu no pai ou na mãe?
 - Sua explicação pode distinguir se o comportamento cromossômico anormal ocorreu na primeira ou na segunda divisão de meiose?
 - Agora suponha que um homem daltônico Klinefelter tem genitores com visão normal, e responda às partes a, b e c.
36. a. Como você sintetiza um pentaplóide?
b. Como você sintetiza um triplóide com o genótipo $A/a/a$?
c. Você obteve uma rara mutação recessiva a^* em uma planta diplóide, cuja análise mendeliana lhe diz que é A/a^* . A partir dessa planta, como você sintetiza um tetraplóide ($4n$) com o genótipo $A/A/a^*/a^*$?
d. Como você sintetiza um tetraplóide com o genótipo $A/a/a/a$?
37. Suponha que você tenha uma linhagem de camundongo com formas citogenéticas distintas do cromossomo 4. A ponta do cromossomo pode ter uma protuberância (chamada de 4^K), ou um satélite (4^S), ou nada (4). Aqui estão esquemas dos três tipos:



Você cruza uma fêmea $4^K/4^S$ com um macho $4/4$ e descobre que a maioria da prole é $4^K/4$ ou $4^S/4$, como esperado. Entretanto, você ocasionalmente descobre alguns tipos raros como os seguintes (todos os outros cromossomos são normais):

- $4^K/4^K/4$
- $4^K/4^S/4$
- 4^K

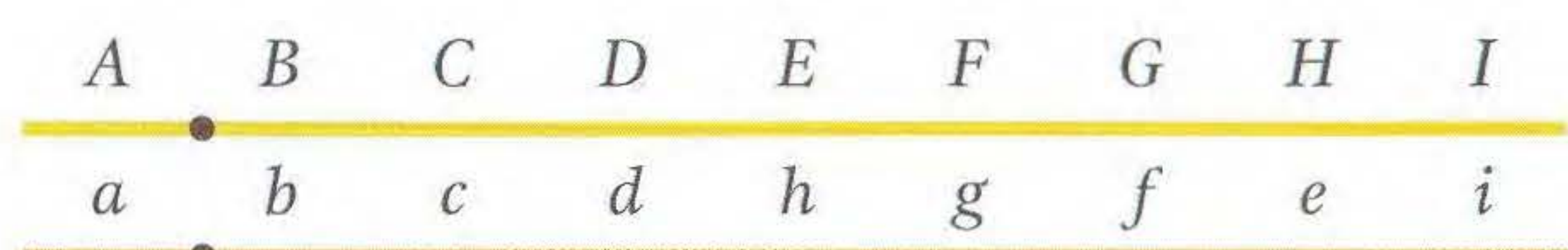
Explique os tipos raros que você encontrou. Cite, o mais precisamente possível, os estágios nos quais eles se originaram, e diga se eles se originaram no genitor masculino, no genitor feminino ou no zigoto. (Cite resumidamente os motivos.)

38. É feito um cruzamento em tomates entre uma planta feminina que é trissômica para o cromossomo 6 e uma planta masculina diplóide normal que é homozigota para o alelo recessivo para folha potato (p/p). Uma planta F_1 trissômica é retrocruzada com o macho de folhas potato.
- Qual a proporção de plantas com folhas normais para plantas com folhas potato quando você supõe que *p* está situado no cromossomo 6?
 - Qual a proporção de folhas normais para folhas potato em plantas quando você supõe que *p* não está situado no cromossomo 6?
39. Uma geneticista de tomates tenta situar cinco mutações recessivas em cromossomos específicos usando trissômicos. Ela cruza cada mutante homozigoto ($2n$) com cada um dos três trissômicos, nos quais os cromossomos 1, 7 e 10 participam. Desses cruzamentos, o geneticista seleciona a prole trissômica (que tem menos vigor) e retrocruza com o recessivo homozigoto apropriado. A prole diplóide desses cruzamentos é examinada. Seus resultados, nos quais as proporções são tipo selvagem:mutante, são as seguintes:

Cromossomo trissômico	Mutação				
	<i>d</i>	<i>y</i>	<i>c</i>	<i>h</i>	<i>cot</i>
1	48:55	72:29	56:50	53:54	32:28
7	52:56	52:48	52:51	58:56	81:40
10	45:42	36:33	28:32	96:50	20:17

Quais das mutações a geneticista pode situar em quais cromossomos? (Explique totalmente sua resposta.)

40. Uma petúnia é heterozigota para os seguintes homólogos autossômicos:



- Desenhe a configuração de pareamento que você veria na metáfase I, e identifique todas as partes de seu diagrama. Numere sequencialmente as cromátides de cima para baixo da página.
 - Ocorre um crossing duplo com três filamentos, com um crossing entre os loci *C* e *D* nas cromátides 1 e 3, e o segundo crossing entre os loci *G* e *H* nas cromátides 2 e 3. Diagrame os resultados desses eventos de recombinação como você veria na anáfase I, e identifique todas as partes de seu diagrama.
 - Desenhe o padrão cromossômico que você veria na anáfase II após os crossings descritos na parte b.
 - Cite os genótipos dos gametas dessa meiose que levariam à formação de uma prole viável. Suponha que todos os gametas são fertilizados pelo pólen que tenha a ordem gênica *A B C D E F G H I*.
41. Dois grupos de geneticistas, na Califórnia e no Chile, começam a trabalhar para desenvolver um mapa de liga-

ção de *medfly*. Eles, independentemente, encontraram que os loci para a cor de corpo (B = preto, b = cinza) e a forma do olho (R = redondo, r = estrela) distam 28 u.m. Eles trocaram linhagens uns com os outros e fizeram cruzamentos; um resumo de todos os seus achados é mostrado aqui:

Cruzamento	F_1	Prole $F_1 \times$ qualquer $b\ r/b\ r$	
$B\ R/B\ R$ (Calif.) $\times\ b\ r/b\ r$ (Calif.)	$B\ R/b\ r$	$B\ R/b\ r$	36%
		$b\ r/b\ r$	36
		$B\ r/b\ r$	14
		$b\ R/b\ r$	14
$B\ R/B\ R$ (Chile) $\times\ b\ r/b\ r$ (Chile)	$B\ R/b\ r$	$B\ R/b\ r$	36
		$b\ r/b\ r$	36
		$B\ r/b\ r$	14
		$b\ R/b\ r$	14
$B\ R/B\ R$ (Calif.) $\times\ b\ r/b\ r$ (Chile)	$B\ R/b\ r$	$B\ R/b\ r$	48
		$b\ r/b\ r$	48
ou		$B\ r/b\ r$	2
$b\ r/b\ r$ (Calif.) $\times\ B\ R/B\ R$ (Chile)		$b\ R/b\ r$	2

- a. Forneça uma hipótese genética que explique os três conjuntos de resultados dos cruzamentos-teste.
- b. Desenhe as características cromossômicas principais da meiose na F_1 de um cruzamento das linhagens californianas e chilenas.

42. Uma planta de milho aberrante dá os seguintes valores de FR quando submetida a cruzamento-teste:

	Intervalo				
	$d-f$	$f-b$	$b-x$	$x-y$	$y-p$
Controle	5	18	23	12	6
Planta aberrante	5	2	2	0	6

(A ordem dos loci é centrômero- $d-f-b-x-y-p$.) A planta aberrante é uma planta saudável, mas ela produz menos ovócitos e pólen que o normal da planta controle.

- a. Proponha uma hipótese para explicar os valores anormais de recombinação e a fertilidade reduzida na planta aberrante.
- b. Use diagramas para explicar a origem dos recombinantes de acordo com sua hipótese.
43. Os seguintes loci de milho estão em um braço do cromossomo 9 na ordem indicada (as distâncias entre eles são mostradas em unidades de mapa):

$c-bz-wx-sh-d$ —centrômero
12 8 10 20 10

C dá aleurona colorida; c , aleurona branca.
 Bz dá folhas verdes; bz , folhas bronze.
 Wx dá sementes com amido; wx , sementes graxas.
 Sh dá sementes lisas; sh , sementes murchas.
 D dá plantas altas; d , anã.

Uma planta de um estoque padrão que é homozigota para todos os cinco alelos recessivos é cruzada com uma planta tipo selvagem do México que é homozigota para todos os cinco alelos dominantes. As plantas F_1 expressam todos os alelos dominantes e, quando retrocruzadas com o genitor recessivo, dão os seguintes fenótipos da prole:

colorida, verde, graxa, lisa, alta	360
branca, bronze, amilácea, murcha, anã	355
colorida, bronze, amilácea, murcha, anã	40
branca, verde, amilácea, lisa, alta	46
colorida, verde, amilácea, lisa, anã	85
branca, bronze, graxa, murcha, alta	84
colorida, bronze, graxa, murcha, alta	8
branca, verde, amilácea, lisa, anã	9
colorida, verde, graxa, lisa, alta	7
branca, bronze, amilácea, murcha, anã	6

Proponha uma hipótese para explicar esses resultados. Inclua

- a. uma explicação geral de sua hipótese, com diagramas se necessário;
- b. por que existem 10 classes;
- c. uma explicação da origem de cada classe, incluindo sua frequência; e
- d. pelo menos um teste de sua hipótese.
44. Plantas de milho cromossomicamente normais têm um locus p no cromossomo 1 e um locus s no cromossomo 5.

P dá folhas verde-escuras; p , folhas verde-claras.
 S dá espigas grandes; s , espigas murchas.

Uma planta original com o genótipo $P/p ; S/s$ tem o fenótipo esperado (verde-escuro, espigas grandes), mas dá resultados inesperados nos cruzamentos, como se segue:

- Na autofecundação, a fertilidade é normal, mas a frequência dos tipos $p/p ; s/s$ é $\frac{1}{4}$ (não $\frac{1}{16}$ como esperado).
- Quando cruzada com um testador normal de genótipo $p/p ; s/s$, a prole F_1 é $\frac{1}{2} P/p ; S/s$ e $\frac{1}{2} p/p ; s/s$; a fertilidade é normal.
- Quando uma planta $F_1\ P/p ; S/s$ é cruzada com uma testadora normal $p/p ; s/s$, ela se mostra semi-estéril, mas, novamente, a prole é $\frac{1}{2} P/p ; S/s$ e $\frac{1}{2} p/p ; s/s$.

Explique esses resultados, mostrando todos os genótipos da planta original, a testadora e as plantas F_1 . Como você testaria sua hipótese?

45. Um rato macho que é fenotipicamente normal apresenta anomalias reprodutivas quando comparado com ratos machos normais, como mostrado no quadro a seguir. Proponha uma explicação genética para esses resultados incomuns, e indique como sua idéia pode ser testada.

Cruzamento	Embriões (número médio)			
	Implantados na parede uterina	Degenerando após implantação	Normais	Degeneração (%)
♂ excepcional × ♀ normal	8,7	5,0	3,7	57,5
♂ normal × ♀ normal	9,5	0,6	8,9	6,5

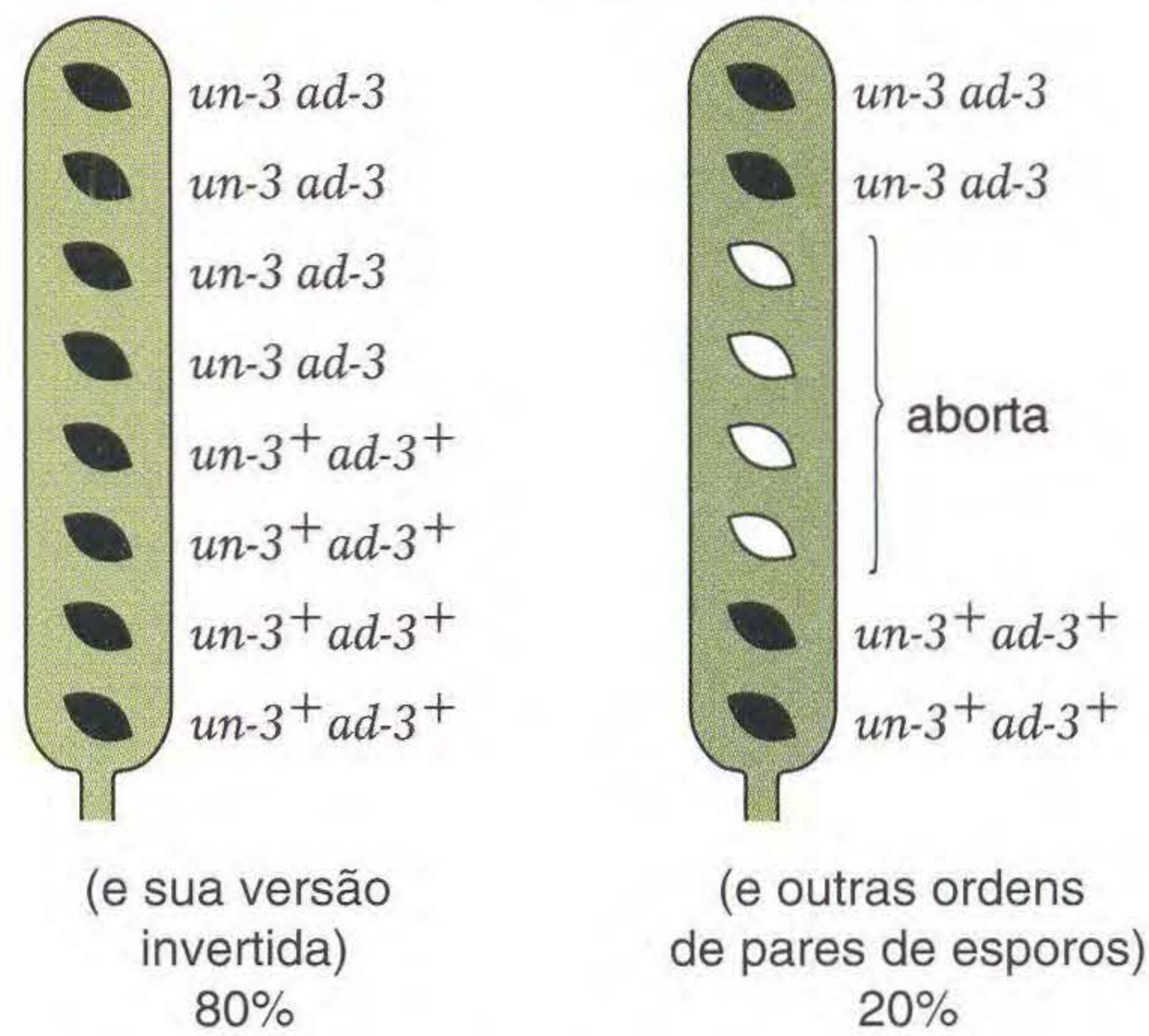
46. Uma geneticista de tomate trabalhando com *Fr*, um alelo mutante dominante que causa o rápido amadurecimento do fruto, decide descobrir qual cromossomo contém esse gene usando um conjunto de linhagens, cada uma das quais é trissômica para um cromossomo. Para fazer isso, ela cruza um mutante diplóide homozigoto com cada uma das linhagens trissômicas tipo selvagem.
- a. Uma planta trissômica F_1 é cruzada com uma planta diplóide tipo selvagem. Qual a proporção de plantas de amadurecimento rápido para lento na prole diplóide desse segundo cruzamento se *Fr* está no cromossomo trissômico? Use diagramas para explicar.
- b. Qual a proporção de plantas de amadurecimento rápido para lento na prole diplóide desse segundo cruzamento se *Fr* não estiver situado no cromossomo trissômico? Use diagramas para explicar.
- c. Aqui estão os resultados dos cruzamentos. Em que cromossomo está *Fr*, e por quê?

Cromossomo trissômico	Amadurecimento rápido:amadurecimento lento em prole diplóide
1	45:47
2	33:34
3	55:52
4	26:30
5	31:32
6	37:41
7	44:79
8	49:53
9	34:34
10	37:39

(O problema 46 é de Tamara Western.)

PROBLEMAS DESAFIADORES

47. O locus *un-3* de *Neurospora* está perto do centrômero do cromossomo 1, e os crossings entre *un-3* e o centrômero são muito raros. O locus *ad-3* está do outro lado do centrômero do mesmo cromossomo, e os crossings ocorrem entre *ad-3* e o centrômero em cerca de 20% das meioses (não ocorrem crossings múltiplos).
- a. Que tipos de ascos lineares (veja o Capítulo 4) você pode prever, e com que frequências, em um cruzamento normal *un-3 ad-3* × tipo selvagem? (Especifique os genótipos dos esporos nos ascos.)
- b. Na maioria das vezes, tais cruzamentos comportam-se previsivelmente, mas, em um caso, uma linhagem padrão *un-3 ad-3* foi cruzada com um tipo selvagem isolado de um canavial no Havaí. Os resultados foram:



Explique esses resultados, e mostre como você pode testar sua idéia. (Nota: Em *Neurospora*, os ascosporos com material cromossômico extra sobrevivem e são de cor preta normal, enquanto os ascosporos sem nenhuma região cromossômica são brancos e inviáveis.)

48. Duas mutações em *Neurospora*, *ad-3* e *pan-2*, estão situadas nos cromossomos 1 e 6, respectivamente. Uma linhagem incomum *ad-3* surge no laboratório, dando os resultados mostrados no quadro a seguir. Explique todos os três resultados com a ajuda de diagramas claramente marcados. (Nota: Em *Neurospora*, os ascosporos com material cromossômico extra sobrevivem e são de cor normal preta, enquanto os ascosporos que não têm nenhuma região cromossômica são brancos e inviáveis.)

	Aspecto do ascosporo	FR entre <i>ad-3</i> e <i>pan-2</i>
1. Normal <i>ad-3</i> × normal <i>pan-2</i>	Todos pretos	50%
2. Anormal <i>ad-3</i> × normal <i>pan-2</i>	Cerca de $\frac{1}{2}$ pretos e $\frac{1}{2}$ brancos (inviáveis)	1%
3. Dos esporos pretos do cruzamento 2, cerca de 50% eram completamente normais, e 50% repetiram o mesmo comportamento que a linhagem original anormal <i>ad-3</i>		

49. Deduza as proporções fenotípicas na prole dos seguintes cruzamentos de autotetraplóides nos quais o locus a^+/a está bem próximo do centrômero. (Suponha que os quatro cromossomos homólogos de qualquer um dos tipos faz par aleatoriamente dois a dois e apenas uma cópia do alelo a^+ é necessária para o fenótipo tipo selvagem.)
- a. $a^+/a^+/a/a \times a/a/a/a$
- b. $a^+/a/a/a \times a/a/a/a$
- c. $a^+/a/a/a \times a^+/a/a/a$
- d. $a^+/a^+/a/a \times a^+/a/a/a$

50. A espécie de algodão do Novo Mundo *Gossypium hirsutum* tem um número $2n$ de cromossomos igual a 52. As espécies do Velho Mundo *G. thurberi* e *G. herbaceum* têm cada uma um número $2n$ igual a 26. Os híbridos entre essas espécies mostram os seguintes arranjos de pareamento cromossômico na meiose:

Híbrido	Arranjo de pareamento
<i>G. hirsutum</i>	13 pequenos bivalentes
× <i>G. thurberi</i>	+ 13 grandes univalentes
<i>G. hirsutum</i>	13 grandes bivalentes
× <i>G. herbaceum</i>	+ 13 pequenos univalentes
<i>G. thurberi</i>	13 grandes univalentes
× <i>G. herbaceum</i>	+ 13 pequenos univalentes

Faça diagramas para interpretar essas observações filogeneticamente, indicando claramente as correlações entre as espécies. Como você faria para provar que sua interpretação está correta? (O problema 50 é adaptado de A.M. Srb, R.D. Owen, and R.S. Edgar, *General Genetics*, 2nd ed. W.H. Freeman and Company, 1965.)

51. Existem seis principais espécies no gênero *Brassica*: *B. carinata*, *B. campestris*, *B. nigra*, *B. oleracea*, *B. juncea* e *B. napus*. Você pode deduzir as relações entre essas seis espécies pelo seguinte quadro:

Espécies ou híbridos F_1	Nº de cromossomos	Nº de bivalentes	Nº de univalentes
<i>B. juncea</i>	36	18	0
<i>B. carinata</i>	34	17	0
<i>B. napus</i>	38	19	0
<i>B. juncea</i> × <i>B. nigra</i>	26	8	10
<i>B. napus</i> × <i>B. campestris</i>	29	10	9
<i>B. carinata</i> × <i>B. oleracea</i>	26	9	8
<i>B. juncea</i> × <i>B. oleracea</i>	27	0	27
<i>B. carinata</i> × <i>B. campestris</i>	27	0	27
<i>B. napus</i> × <i>B. nigra</i>	27	0	27

- a. Deduza o número de cromossomos de *B. campestris*, *B. nigra* e *B. oleracea*.
b. Mostre claramente qualquer correlação evolutiva entre as seis espécies que você pode deduzir a nível cromossômico.

52. Vários tipos de mosaicismo sexual estão bem documentados em humanos. Sugira como cada um dos seguintes exemplos pode ter surgido por não-disjunção na mitose:
a. XX/XO (isto é, existem dois tipos de células no corpo, XX e XO)
b. XX/XXYY
c. XO/XXX

- d. XX/XY
e. XO/XX/XXX

53. Em *Drosophila*, um cruzamento (cruzamento 1) foi feito entre duas moscas mutantes, uma homozigota para a mutação recessiva asa inclinada (*b*) e a outra homozigota para a mutação recessiva sem olhos (*e*). As mutações *e* e *b* são alelos de dois genes diferentes que são conhecidos como estando proximamente ligados no pequeno cromossomo autossômico 4. Toda a prole tinha um fenótipo tipo selvagem. Uma da prole de fêmeas foi cruzada com um macho de genótipo *b e/b e*; chamaremos esse cruzamento de 2. A maior parte da prole do cruzamento 2 era dos tipos esperados, mas também existiam raras fêmeas com o fenótipo tipo selvagem.

a. Explique qual seria a prole comumente esperada do cruzamento 2.

b. Poderiam as fêmeas da forma rara tipo selvagem ter surgido por (1) crossing over ou (2) não-disjunção? Explique.

c. As fêmeas raras tipo selvagem foram submetidas a um cruzamento-teste com um macho de genótipo *b e/b e* (cruzamento 3). A prole era

$\frac{1}{6}$ tipo selvagem

$\frac{1}{6}$ curva, sem olhos

$\frac{1}{3}$ curva

$\frac{1}{3}$ sem olhos

Qual das explicações na parte b é compatível com esse resultado? Explique os genótipos e fenótipos da prole do cruzamento 3 e suas proporções.

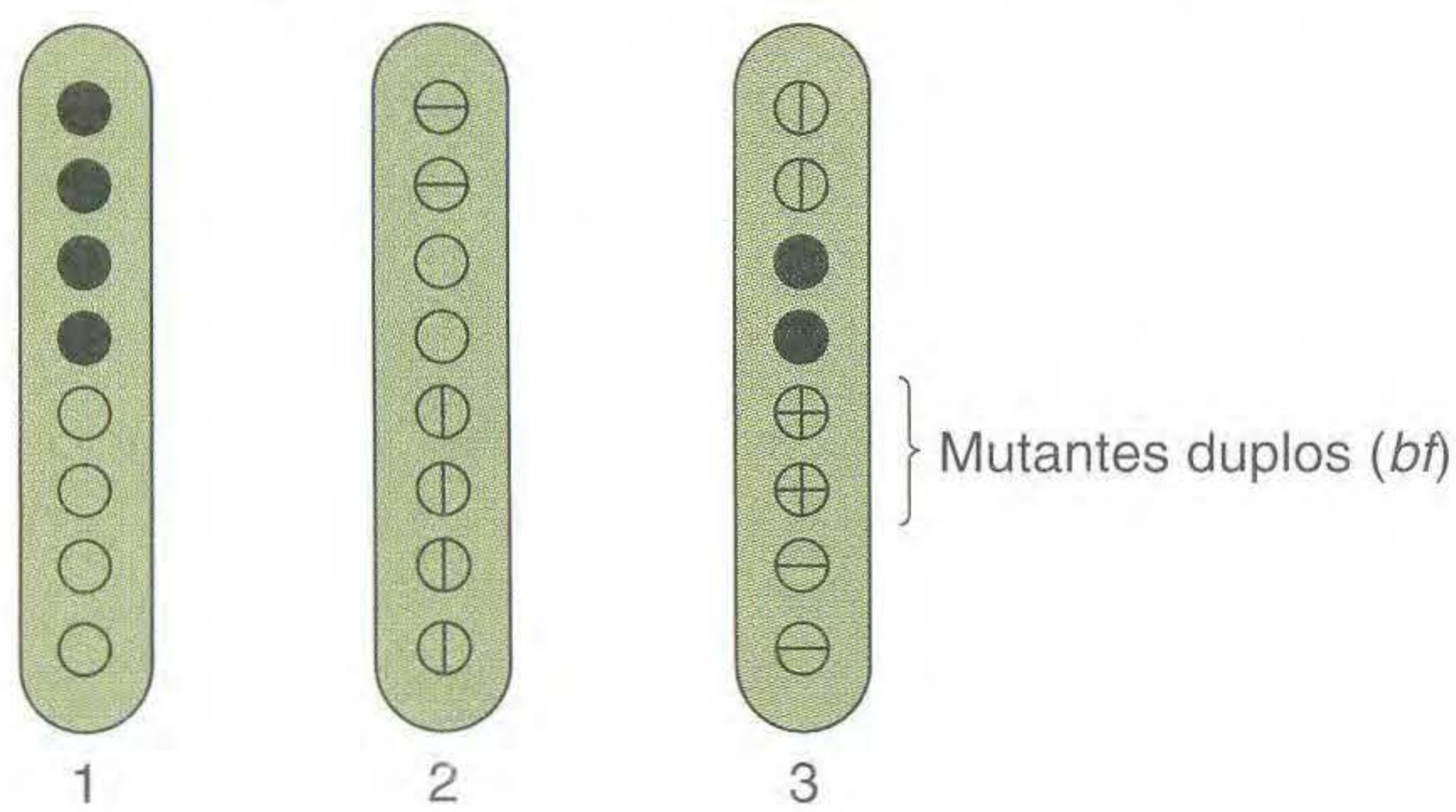


Solucionando o Problema 53

- Defina homozigoto, mutação, alelo, proximamente ligados, recessivo, tipo selvagem, crossing over, não-disjunção, cruzamento-teste, fenótipo e genótipo.
- Este problema tem a ver com ligação ao sexo? Explique.
- Quantos cromossomos as drosófilas têm?
- Desenhe um heredograma claro resumindo os resultados dos cruzamentos 1, 2 e 3.
- Desenhe os gametas produzidos por ambos os genitores no cruzamento 1.
- Desenhe a constituição do cromossomo 4 da prole do cruzamento 1.
- É surpreendente que a prole do cruzamento 1 seja de fenótipo tipo selvagem? O que esse resultado lhe diz?
- Desenhe a constituição do cromossomo 4 do macho testador no cruzamento 2 e os gametas que ele pode produzir.
- Com relação ao cromossomo 4, que gametas podem produzir a genitora no cruzamento 2, na ausência de não-disjunção? Quais seriam comuns e quais seriam raros?

10. Desenhe a não-disjunção meiótica da primeira e segunda divisões no genitor feminino do cruzamento 2, bem como nos gametas resultantes.
11. Algum dos gametas da parte 10 é aneuplóide?
12. Você esperaria que gametas aneuplóides dessem origem a uma prole viável? Essa prole seria nulissômica, monossômica, dissômica ou trissômica?
13. Que fenótipos da prole seriam produzidos pelos vários gametas considerados nas partes 9 e 10?
14. Considere a proporção fenotípica na prole do cruzamento 3. Muitas proporções genéticas são baseadas em metade e um quarto, mas essa proporção é baseada em terços e sextos. O que essas proporções indicam?
15. Pode ter algum significado o fato de que os cruzamentos envolvem genes em um cromossomo muito pequeno? Quando o tamanho do cromossomo é relevante em genética?
16. Desenhe a prole esperada do cruzamento 3 conforme as duas hipóteses, e cite alguma idéia sobre proporções relativas.

54. No fungo *Ascobolus* (similar a *Neurospora*), os ascosporos são normalmente pretos. A mutação *f*, que produz ascosporos de cor castanho-clara, está em um gene à direita do centrômero do cromossomo 6, enquanto a mutação *b*, que produz ascosporos beges, está em um gene à esquerda do mesmo cromossomo. Em um cruzamento de genitores castanho-claro e bege ($+f \times b+$), a maioria das óctades mostraram quatro ascosporos castanhos e quatro ascosporos beges, mas três raras óctades excepcionais foram encontradas, como mostrado na ilustração seguinte. No desenho, preto é o fenótipo tipo selvagem, uma linha vertical é castanha, uma linha horizontal é bege e um círculo vazado representa um ascosporo abortado (morto).



- a. Cite explicações razoáveis para essas três óctades excepcionais.
- b. Diagrame a meiose que deu origem à óctade 2.

55. O ciclo de vida do fungo haplóide *Ascobolus* é similar ao de *Neurospora*. Um tratamento mutacional produziu duas linhagens mutantes, 1 e 2, ambas as quais, quando cruzadas com tipo selvagem, dão tétrades desordenadas, todas dos seguintes tipos (bege é a cor castanho-clara; normalmente, os cruzamentos produzem ascosporos todos pretos):

par de esporos 1	preto	par de esporos 3	castanho
par de esporos 2	preto	par de esporos 4	castanho

- a. O que esse resultado mostra? Explique.

As duas linhagens mutantes foram cruzadas. A maioria das tétrades desordenadas eram do seguinte tipo:

par de esporos 1	castanho	par de esporos 3	castanho
par de esporos 2	castanho	par de esporos 4	castanho

- b. O que esse resultado sugere? Explique.

Quando grandes números de tétrades desordenadas foram triadas ao microscópio, alguma raras que continham esporos pretos foram encontradas. Quatro casos são mostrados aqui:

	Caso A	Caso B	Caso C	Caso D
par de esporos 1	pretos	pretos	pretos	pretos
par de esporos 2	pretos	castanhos	pretos	abortam
par de esporos 3	castanhos	castanhos	abortam	castanhos
par de esporos 4	castanhos	castanhos	abortam	castanhos

(Nota: Ascosporos com material genético extra sobrevivem, mas aqueles com menos de um genoma haplóide abortam.)

- c. Proponha explicações genéticas razoáveis para cada um desses quatro casos raros.
- d. Você acha que mutações nas duas linhagens mutantes originais estavam em um só gene? Explique.

Genética de Populações

17



Polimorfismo de cor de conchas em *Liguus fasciatus*. [De David Hillis. *J. Hered.*, July-August 1991.]

A té agora, em nossa investigação sobre genética, nos concentramos em processos que ocorrem em organismos individuais e células. Como a célula copia o DNA, e quais as causas das mutações? Como os mecanismos de segregação e recombinação afetam os tipos e proporções de gametas produzidos por um organismo individual? Como o desenvolvimento de um organismo é afetado pelas interações entre seu DNA, a maquinaria celular de síntese de proteínas, os processos metabólicos celulares e o ambiente externo? Mas os organismos não vivem apenas como indivíduos isolados. Eles interagem uns com os outros em grupos, **populações**, e existem questões sobre a composição genética dessas populações que não podem ser respondidas apenas com o conhecimento dos processos genéticos básicos individualmente. Por que os alelos dos genes que codificam fatores proteicos VIII e IX que causam uma falha da coagulação normal, a hemofilia, são tão raros em todas as populações

Perguntas Fundamentais

- Quanto de variação genética há nas populações naturais de organismos?
- Quais são os efeitos dos padrões de reprodução na variação genética?
- Quais são as fontes de variação genética que são observadas nas populações?
- Quais são os processos que causam mudanças no tipo e quantidade de variação genética nas populações?

Tópicos

- 17.1 Variação e sua modulação
- 17.2 Efeitos da reprodução sexual na variação
- 17.3 Fontes de variação
- 17.4 Seleção
- 17.5 Polimorfismo balanceado
- 17.6 Eventos aleatórios

humanas? Em contraste, um alelo, Hb^S , do gene de hemoglobina β que causa a anemia falciforme é muito comum em partes da África e do Mediterrâneo. O que responde pelas diferenças de frequências dessas duas anomalias sangüíneas humanas? Que mudanças na frequência de anemia falciforme são esperadas nos descendentes de africanos na América do Norte como consequência de mudança no ambiente e de entrecruzamento de africanos e nativos americanos? Essas são perguntas sobre o que determina a composição genética das populações e como se espera que essa composição possa mudar com o tempo. Essas perguntas constituem o domínio da **genética de populações**.

A composição genética de uma população é a coleção de frequências de genótipos diferentes nessa população. Essas frequências são consequência dos processos que atuam em organismos individuais para aumentar ou diminuir o número de organismos de cada genótipo. Para relacionar os processos genéticos básicos em âmbito individual à composição genética da população, devemos investigar os seguintes fenômenos:

1. O efeito dos *padrões de reprodução* em genótipos diferentes na população. Os membros individuais de uma população podem se reproduzir aleatoriamente ou podem se reproduzir preferencialmente com parentes próximos (*endogamia*) ou preferencialmente com indivíduos de genótipos ou fenótipos similares ou diferentes (*reprodução aleatória*).
2. As mudanças na composição da população devidas à *migração* de indivíduos entre populações.
3. A taxa de introdução de novas variações genéticas na população por *mutação*, que traz novos alelos nos loci.
4. A produção de novas combinações de características por *recombinação*, que redistribuem combinações de alelos em loci diferentes.
5. As mudanças na composição da população devidas ao efeito da *seleção natural*. Genótipos diferentes podem ter taxas diferentes de reprodução, e prole geneticamente diferente pode ter chances diferenciais de sobrevivência.
6. As consequências das *flutuações aleatórias* nas taxas reprodutivas de genótipos diferentes. Como qualquer determinado indivíduo tem apenas alguns filhos e o tamanho total da população é limitado, as proporções genéticas das meioses nunca são exatamente como previstas pela teoria em famílias reais e populações reais. Essa flutuação aleatória causa *deriva genética* nas frequências alélicas de geração a geração.

► 17.1 Variação e Sua Modulação

A genética de populações é tanto uma ciência experimental quanto teórica. Do lado experimental, ela fornece descrições dos padrões reais de variação genética entre membros individuais das populações e estima as taxas dos processos de reprodução, mutação, recombinação e seleção natural. Também estima a variação aleatória nas taxas reprodutivas. Do lado teórico, faz previsões sobre como se espera que a composição genética das populações possa mudar como consequência das várias forças que nelas operam.

Observações da variação

Os estudos de genética de populações foram capazes de explorar apenas conjuntos limitados de características, devido à necessidade de uma relação simples entre a variação fenotípica e a genotípica. A variação entre fenótipo e genótipo varia em simplicidade, dependendo da característica que é observada. Em um extremo, o fenótipo de interesse pode ser o mRNA codificado por um trecho do genoma. No outro extremo está o grosso das características de interesse para os criadores de animais e plantas e para a maioria dos evolucionistas: as variações em produtividade, taxa de crescimento, forma do corpo, taxa metabólica e comportamento, que constituem as diferenças óbvias entre variedades e espécies. Essas características têm uma relação muito complexa com o genótipo. Não há alelo para ter uma altura específica. Tais diferenças, sendo consequência de variação genética, serão afetadas por vários genes e também pela variação ambiental. Para suas análises, devemos usar os métodos introduzidos no Capítulo 18 para dizer algo sobre os genótipos subjacentes à variação fenotípica. Mas, como veremos no Capítulo 18, não é possível fazer afirmativas precisas sobre a variação genotípica subjacente a tais características. Por esse motivo, a maioria dos estudos experimentais de genética de populações concentrou-se em características com relações simples com o genótipo. Os fenótipos diferentes para uma característica podem ser mostrados como resultando de formas alélicas diferentes de um único gene. Um objeto favorito para o estudo de geneticistas de populações humanas, por exemplo, têm sido os vários grupos sangüíneos humanos. O fenótipo de um grupo sangüíneo é a presença de determinados antígenos na superfície das hemácias de determinados anticorpos no soro sangüíneo. Os fenótipos qualitativamente distintos de determinado grupo sangüíneo, digamos, grupo MN, são codificados por alelos alternativos em um único locus, e os fenótipos são insensíveis a variações ambientais. Assim, a variação observada nos tipos sangüíneos é totalmente consequência de simples diferenças genéticas.

O estudo da variação consiste em dois estágios. O primeiro é uma descrição da variação fenotípica. O segundo é a tradução desses fenótipos em termos genéticos e uma descrição da variação genética subjacente. Se houver uma perfeita correspondência de um para um entre genótipo e fenótipo, então essas duas etapas fundem-se em uma, como no grupo sangüíneo MN. Se a relação é mais complexa, por exemplo, como resultado de dominância, os heterozigotos assemelham-se aos homozigotos, e pode ser necessário fazer cruzamentos experimentais ou observar heredogramas para traduzir fenótipos em genótipos. Esse é o caso para outro grupo sangüíneo humano, o sistema ABO, para o qual existem dois alelos dominantes, I^A e I^B , e um alelo recessivo i . Os indivíduos com o tipo sangüíneo A ou B podem ser homozigotos para seus respectivos alelos ($I^A I^A$ ou $I^B I^B$) ou heterozigotos para seu tipo de alelo e o alelo recessivo ($I^A i$ ou $I^B i$).

A descrição mais simples da variação de um único gene é a lista de proporções observadas de genótipos em uma população. Tais proporções são chamadas de **frequências genotípicas**. O Quadro 17.1 mostra essa distribuição de fre-

Quadro 17.1 Freqüências Genotípicas para Alelos no Locus de Grupo Sangüíneo MN em Várias Populações Humanas

População	Genótipo			Freqüências alélicas	
	M/M	M/N	N/N	p(M)	q(N)
Esquimós	0,835	0,156	0,009	0,913	0,087
Aborígenes australianos	0,024	0,304	0,672	0,176	0,824
Egípcios	0,278	0,489	0,233	0,523	0,477
Alemães	0,297	0,507	0,196	0,550	0,450
Chineses	0,332	0,486	0,182	0,575	0,425
Nigerianos	0,301	0,495	0,204	0,548	0,452

Fonte: W.C. Boyd, *Genetics and the Races of Man*. D.C. Heath, 1950.

qüência para os três genótipos do gene MN em várias populações humanas. Note que há uma variação entre membros individuais de cada população, pois existem genótipos diferentes presentes, e há uma variação nas freqüências desses genótipos de população em população. Por exemplo, a maioria das pessoas na população de esquimós são MM, enquanto esse genótipo é muito raro entre aborígenes australianos.

Mais tipicamente, em vez das freqüências de genótipos diplóides, são usadas as freqüências de alelos alternativos. A freqüência de alelos é simplesmente a proporção dessa forma alélica do gene entre todas as cópias do gene na população, onde cada organismo diplóide na população é contado como contribuindo com dois alelos de cada gene. Os homozigotos para um alelo têm duas cópias desse alelo, enquanto os heterozigotos têm apenas uma cópia. Assim, a freqüência de um alelo é a freqüência de homozigotos mais metade da freqüência de heterozigotos. Assim, se a freqüência de indivíduos A/A fosse, digamos, 0,36 e a freqüência de indivíduos A/a fosse 0,48, a freqüência do alelo A seria $0,36 + 0,48/2 = 0,60$. O Boxe 17.1 dá a forma geral desse cálculo. O Quadro 17.1 mostra os valores de *p* e *q*, a **freqüência alélica** dos dois alelos M e N do grupo sangüíneo MN em populações diferentes.

A variação simples pode ser observada dentro e dentre populações de qualquer espécie em vários níveis, do fenóti-

po da morfologia externa até as seqüências de aminoácidos das proteínas. De fato, a variação genotípica pode ser diretamente caracterizada pelo seqüenciamento do DNA para o mesmo gene ou para as mesmas regiões entre genes de muitos indivíduos. Cada espécie de organismo examinado revelou uma considerável variação genética, ou **polimorfismo**, manifesta como um ou mais níveis de observação dentro das populações e/ou dentre as populações. Um gene ou uma característica fenotípica é dita como sendo polimórfica se houver mais de uma forma do gene ou mais de um fenótipo para essa característica em uma população. Em alguns casos, quase toda a população é caracterizada por uma forma do gene ou característica, com raros indivíduos excepcionais levando uma variante incomum. Essa forma extremamente comum é chamada de **tipo selvagem**, em contraste com os mutantes raros. Em outros casos, são comuns duas ou mais formas, e não é possível identificar uma que é o tipo selvagem. A variação genética que pode ser a base da mudança evolutiva é ubíqua.

É impossível neste livro dar um quadro adequado da imensa riqueza ou mesmo a simples variação genética que existe nas espécies. Podemos considerar apenas alguns exemplos de tipos diferentes para ter uma noção da diversidade genética dentro de uma espécie. Cada um desses exemplos pode ser multiplicado várias vezes em outras espécies e com outras características.

Boxe 17.1 Cálculo de Freqüências Alélicas

Se $f_{A/A}$, $f_{A/a}$ e $f_{a/a}$ são as freqüências de três genótipos em um locus com dois alelos, então a freqüência *p* do alelo A e a freqüência *q* do alelo *a* são obtidas contando alelos. Como cada homozigoto A/A consiste apenas em alelos A e como metade dos alelos de cada heterozigoto A/a são alelos A, a freqüência total *p* dos alelos A na população é calculada como

$$p = f_{A/A} + \frac{1}{2} f_{A/a} = \text{freqüência de A}$$

Similarmente, a freqüência *q* do alelo *a* é dada por

$$q = f_{a/a} + \frac{1}{2} f_{A/a} = \text{freqüência de a}$$

Portanto,

$$p + q = f_{A/A} + f_{a/a} + f_{A/a} = 1,00$$

e

$$q = 1 - p$$

Se existirem mais de duas formas alélicas diferentes, a freqüência de cada alelo é simplesmente a freqüência de seu homozigoto mais metade da soma das freqüências de todos os heterozigotos nos quais aparecem.

Polimorfismos de proteínas

Polimorfismos imunológicos Vários loci em vertebrados codificam especificidades antigênicas, tais como os tipos sanguíneos ABO. São conhecidas mais de 40 especificidades diferentes para antígenos de hemácias humanas, e várias centenas são conhecidas em gado. Outro importante polimorfismo em humanos é o sistema HLA de antígenos celulares, que estão implicados na compatibilidade de enxerto de tecidos. O Quadro 17.2 dá as frequências alélicas do locus de grupo sanguíneo ABO em algumas populações muito diferentes. O polimorfismo para o sistema HLA é muito maior. Parece haver dois loci principais, cada um com cinco alelos distinguíveis. Assim, existem $5^2 = 25$ tipos gaméticos diferentes possíveis, dando 25 formas homozigotas diferentes. Existem $n(n - 1)/2$ modos de combinar n coisas diferentes de uma vez, e, portanto, existem $(25)(24)/2 = 300$ heterozigotos diferentes. Entretanto, nem todos os genótipos são fenotipicamente distinguíveis; apenas 121 classes fenotípicas podem ser vistas. Em um estudo de uma amostra apenas de 100 europeus, 53 dos 121 fenótipos possíveis foram de fato observados.

Polimorfismos de seqüências de aminoácidos Os estudos dos polimorfismos genéticos foram feitos ao nível dos polipeptídeos codificados pelas regiões codificantes dos próprios genes. Se há uma mudança não-sinônima de códon em um gene (digamos GGU para GAU), o resultado é uma substituição de aminoácidos no polipeptídeo produzido na tradução (nesse caso, o ácido aspártico é substituído por glicina). A variação na seqüência de aminoácidos de uma proteína pode ser detectada pelo seqüenciamento do DNA que codifica a proteína de um grande número de indivíduos. Esse método seria útil se alguém desejasse conhecer exatamente quais aminoácidos na seqüência de proteínas estão variando, mas é muito demorado e caro fazer tais projetos de seqüenciamento de DNA para muitos genes codificantes de proteínas. Existe, entretanto, um substituto prático para o seqüenciamento do DNA que pode ser usado se houver interesse apenas em detectar formas variantes de uma proteína sem saber quais são as mudanças de aminoácidos. Esse método alternativo usa eletroforese em gel, descrito no Capítulo 1, para detectar proteínas variantes. Esse método usa a mudança nas propriedades físicas de uma proteína quando um aminoácido é substituído por outro. As proteínas têm uma carga final que resulta da ionização das cadeias laterais de cinco aminoácidos (ácido glutâmico, ácido aspártico, arginina, lisina

Quadro 17.2 Frequências Alélicas I^A , I^B e i no Locus do Grupo Sanguíneo ABO em Várias Populações Humanas

População	I^A	I^B	i
Esquimós	0,333	0,026	0,641
Sioux	0,035	0,010	0,955
Belgas	0,257	0,058	0,684
Japoneses	0,279	0,172	0,549
Pigmeus	0,227	0,219	0,554

Fonte: W.C. Boyd, *Genetics and the Races of Man*. D.C. Heath, 1950.

e histidina). As substituições de aminoácidos podem substituir diretamente um desses aminoácidos, ou uma substituição não-carregada perto de um aminoácido carregado pode alterar o grau de ionização do aminoácido carregado, ou uma substituição na junção entre duas hélices α pode causar uma leve mudança na estrutura tridimensional do polipeptídeo dobrado. Em todos esses casos, a carga final do polipeptídeo será alterada.

A eletroforese em gel detecta a mudança na carga final da proteína. A Figura 17.1 mostra o resultado de tal separação eletroforética. As colunas representam variantes de uma enzima esterase em *Drosophila pseudoobscura*, onde cada coluna é a proteína de uma mosca diferente. A Figura 17.2 mostra um gel similar para hemoglobinas humanas variantes. Nesse caso, a maioria dos indivíduos são heterozigotos para uma hemoglobina A variante e uma normal. O Quadro 17.3 mostra as frequências de alelos diferentes para três genes codificantes de enzimas em *D. pseudoobscura* e várias populações: um locus quase polimórfico (desidrogenase málica), um locus moderadamente polimórfico (α -amilase) e um locus altamente polimórfico (xantina desidrogenase).

A técnica de proteínas em gel (bem como o seqüenciamento do DNA) difere fundamentalmente de outros métodos de análise genética, permitindo o estudo de genes que não estão variando em uma população, porque a presença de uma proteína é evidente, por si, da seqüência de DNA codificando a seqüência da proteína. Assim, foi possível indagar que proporção de todos os genes estruturais no genoma de uma espécie é polimórfica e que fração média do genoma do indivíduo está no estado heterozigoto — a *heterozigose* — em uma população. Um grande número de espécies foi amos-

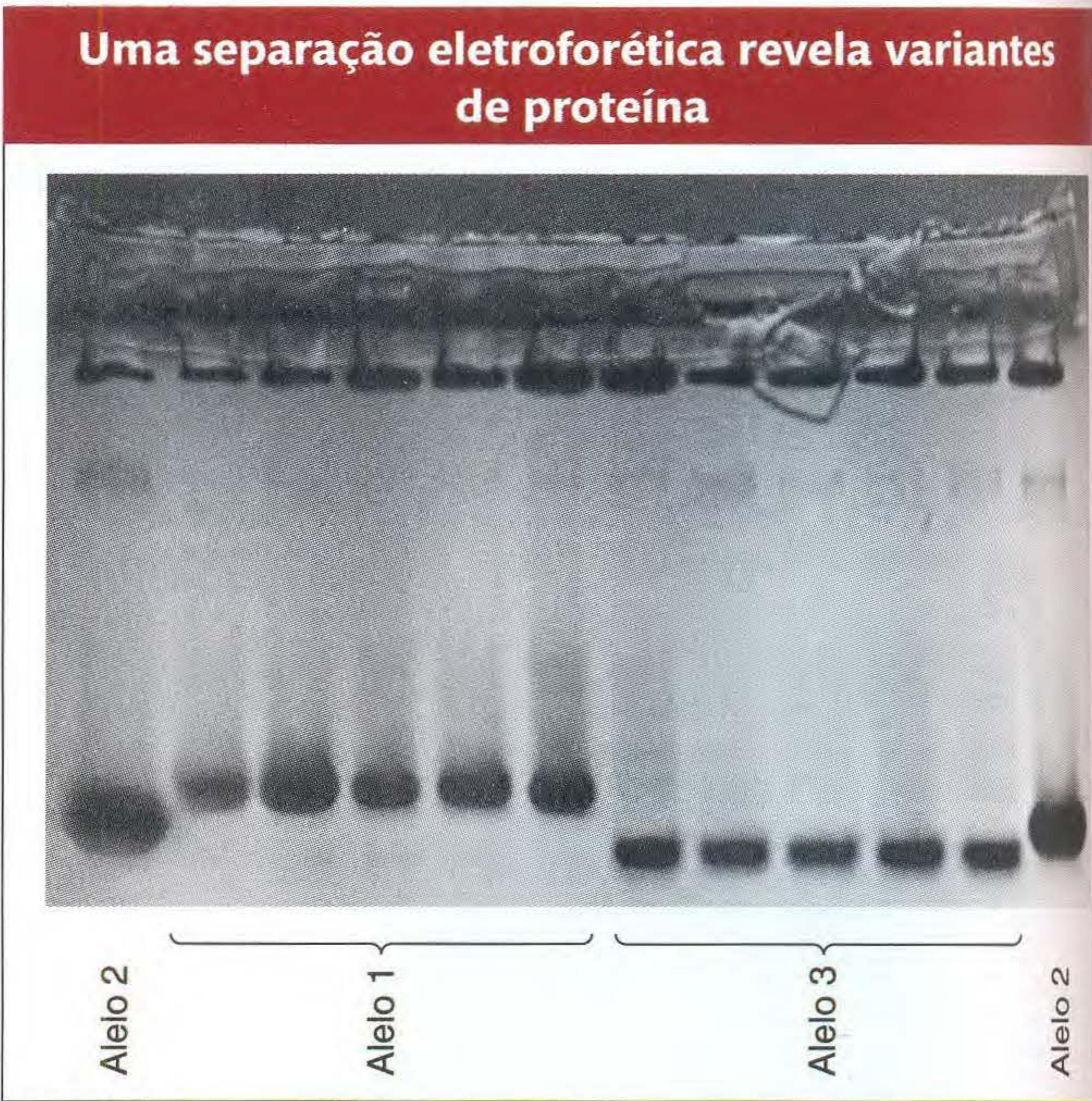


FIG. 17.1 Gel de eletroforese de proteínas codificadas por homozigotos para três alelos diferentes no locus de *esterase-5* em *Drosophila pseudoobscura*. Cada coluna é uma proteína de uma mosca diferente. Amostras repetidas de proteínas codificadas pelo mesmo alelo são idênticas, mas existem diferenças repetíveis entre os alelos.



FIG. 17.2 Gel de eletroforese mostrando a hemoglobina normal A e vários alelos variantes de hemoglobina. Cada coluna representa um indivíduo diferente. Uma das bandas de coloração escura é marcada como hemoglobina A normal. A outra banda de coloração escura vista na maioria das colunas (mais claramente nas colunas 3 e 4) representa qualquer das várias hemoglobinas variantes derivadas do segundo alelo de um heterozigoto. A hemoglobina A está faltando nas colunas 9 e 10 porque os indivíduos são homozigotos para um alelo variante. [Richard C. Lewontin.]

trado por esse método, incluindo bactérias, fungos, plantas e vertebrados superiores e invertebrados. Os resultados são marcadamente consistentes entre as espécies. Cerca de 33% dos genes que codificam proteínas são detectavelmente polimórficos a nível de proteínas, e a heterozigose média em uma população em relação a todos os loci amostrados é de cerca de 10%. Isso significa que a varredura do genoma em praticamente qualquer espécie mostraria que cerca de 1 em cada 10 genes em um indivíduo está em condição heterozigota para as variações genéticas que resultam em diferenças na sequência de aminoácidos das proteínas, e cerca de 33% de todos os genes têm dois ou mais de tais alelos segregando em qualquer população. Assim, o potencial de variação para a evolução é imenso. A desvantagem da técnica de eletroforese é que ela detecta variação apenas nas regiões codificantes de proteínas dos genes, perdendo mudanças importantes em elementos reguladores que são subjacentes à evolução da forma e função.

Estrutura do DNA e polimorfismo de seqüências

A análise do DNA possibilita examinar a variação na estrutura do genoma entre indivíduos e entre espécies. Existem três níveis nos quais tais estudos podem ser feitos. Examinando a variação no número de cromossomos e morfologia, tem-se uma visão em larga escala de reorganizações do genoma. O estudo da variação em sítios reconhecidos por enzimas de restrição dá uma visão geral da variação de pares de bases. Em um nível mais fino, os métodos de seqüenciamento do DNA permitem que a variação seja observada a cada par de bases.

Polimorfismo cromossômico Embora o cariótipo seja geralmente visto como uma característica distinta de uma espécie, de fato várias espécies são polimórficas para o número de cromossomos e morfologia. Cromossomos extras

Quadro 17.3 Freqüências de Vários Alelos em Três Loci Codificantes de Enzimas em Quatro Populações de *Drosophila pseudoobscura*

		População			
Locus (codificante de enzima)	Alelo	Berkeley	Mesa Verde	Austin	Bogotá
Desidrogenase málica	A	0,969	0,948	0,957	1,00
	B	0,031	0,052	0,043	0,00
α-Amilase	A	0,030	0,000	0,000	0,00
	B	0,290	0,211	0,125	1,00
	C	0,680	0,789	0,875	0,00
Xantina desidrogenase	A	0,053	0,016	0,018	0,00
	B	0,074	0,073	0,036	0,00
	C	0,263	0,300	0,232	0,00
	D	0,600	0,581	0,661	1,00
	E	0,010	0,030	0,053	0,00

Fonte: R.C. Lewontin, *The Genetic Basis of Evolutionary Change*. Columbia University Press, 1974.

(supranumerários), translocações recíprocas e inversões são observados em muitas populações de plantas, insetos e mesmo mamíferos. A Figura 17.3 mostra uma variedade de alças de inversão encontradas em populações naturais de *D. pseudoobscura*. Cada uma dessas alças resulta de levar dois cromossomos homólogos, um dos quais contém um trecho que está em ordem linear reversa com relação ao outro cromossomo (veja o Capítulo 16). O Quadro 17.4 dá as frequências de cromossomos supranumerários e translocações heterozigotas em uma população da planta *Clarkia elegans* da Califórnia. O cariótipo “típico” da espécie seria difícil de identificar nessa planta, na qual apenas 56% das plantas individuais não têm cromossomos supranumerários e translocações.

Varição de sítio de restrição Um modo rápido e não-carro de observar os níveis gerais nas seqüências de DNA é digerir esse DNA usando enzimas de restrição (veja o Capítulo 20). Existem muitas enzimas de restrição diferentes, cada uma das quais irá reconhecer uma seqüência de bases diferentes e cortar o DNA no local dessa seqüência. O resultado será dois fragmentos de DNA cujos comprimentos são determinados pelo local do sítio de restrição na molécula original não-cortada. Uma enzima de restrição que reconhece seqüências de seis bases (uma “cortadora de seis”) irá reconhecer uma seqüência apropriada aproximadamente uma vez a cada $4^6 = 4.096$ pares de bases ao longo da molécula de DNA (determinado pela probabilidade de que uma base específica [da qual existem quatro] seja encontrada em cada uma de seis posições). Se houver polimorfismo na população para uma das seis bases do sítio de reconhecimento, então a enzima reconhecerá e cortará o DNA em uma variante e não na outra (veja o Capítulo 20). Assim, haverá um polimorfismo de comprimento de fragmento de restrição (RFLP) na população. Um painel de, digamos, oito enzimas “cortadoras de seis” amostrará cada $4.096/8 = 512$ pares de bases para cada um desses polimorfismos. Entretanto, quando um é encontrado, não sabemos qual dos seis pares de bases no sítio de reconhecimento é polimórfico.

Se usarmos enzimas que reconhecem seqüências de quatro bases (“cortadoras de quatro”), haverá um sítio de reconhecimento a cada $4^4 = 256$ pares de bases, e, assim, um painel de oito enzimas diferentes amostrará cerca de uma vez a cada $256/8 = 32$ pares de bases ao longo da seqüência de DNA. Em adição a mudanças de um par de bases que destroem os sítios de reconhecimento das enzimas de restrição, as inserções ou deleções de trechos do DNA ao longo do filamento de DNA entre os locais dos sítios de restrição farão com que os fragmentos dos tamanhos de restrição variem.

Uma variedade de estudos de diferentes enzimas de restrição foram feitos para regiões diferentes do cromossomo

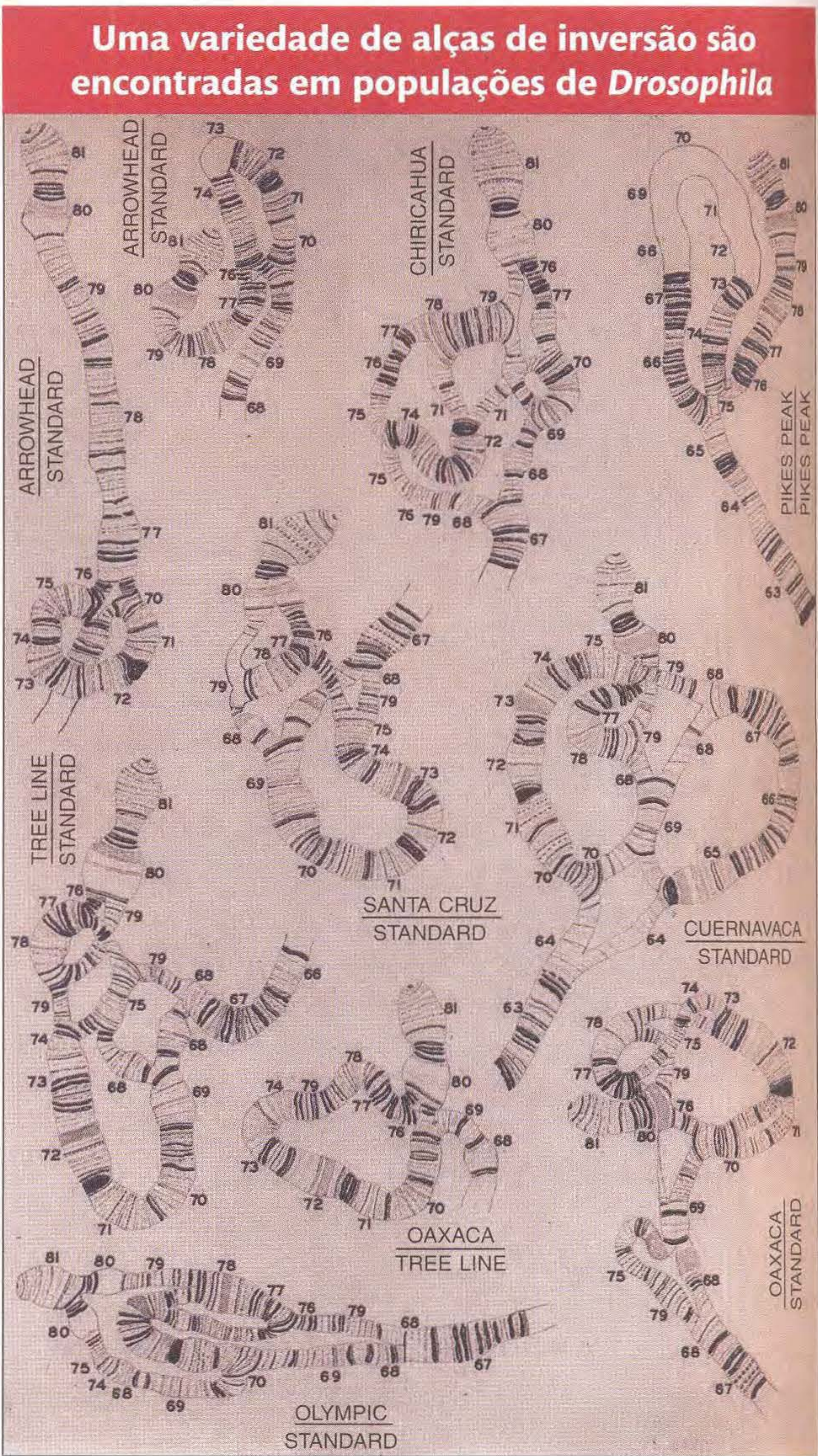


FIG. 17.3 Polimorfismo de inversão do cromossomo 3 em populações naturais de *Drosophila pseudoobscura*. As inversões são denominadas pela localidade na qual foram primeiro observadas. Os pareamentos de diferentes ordens de genes mostram alças nos cromossomos politênicos, revelando a localização dos pontos de quebra das inversões. [T. Dobzhansky. *Chromosomal Races in Drosophila pseudoobscura and Drosophila persimilis*. pp. 47-144. Carnegie Institution of Washington, 1944.]

X e os dois autossomos grandes de *Drosophila melanogaster*. Esses estudos encontraram entre 0,1 e 1,0% de heterozigose por sítio de nucleotídeo, com uma média de 0,4%. O resultado de tal estudo no gene de xantina desidrogenase em *D. pseudoobscura* é mostrado na Figura 17.4. Essa ilustração mostra, simbolicamente, o padrão de restrição de 58 amostras de cromossomos da natureza, polimórficos em 78 sítios de

Quadro 17.4 Frequências de Plantas com Cromossomos Supranumerários e de Translocações Heterozigotas em uma População de <i>Clarkia elegans</i> da Califórnia			
Sem supranumerários ou translocações	Translocações	Supranumerários	Tanto translocações quanto supranumerários
0,560	0,133	0,265	0,042

Fonte: H. Lewis, *Evolution* 5, 1951, 142-157.

Padrões de restrição revelam níveis gerais de variação

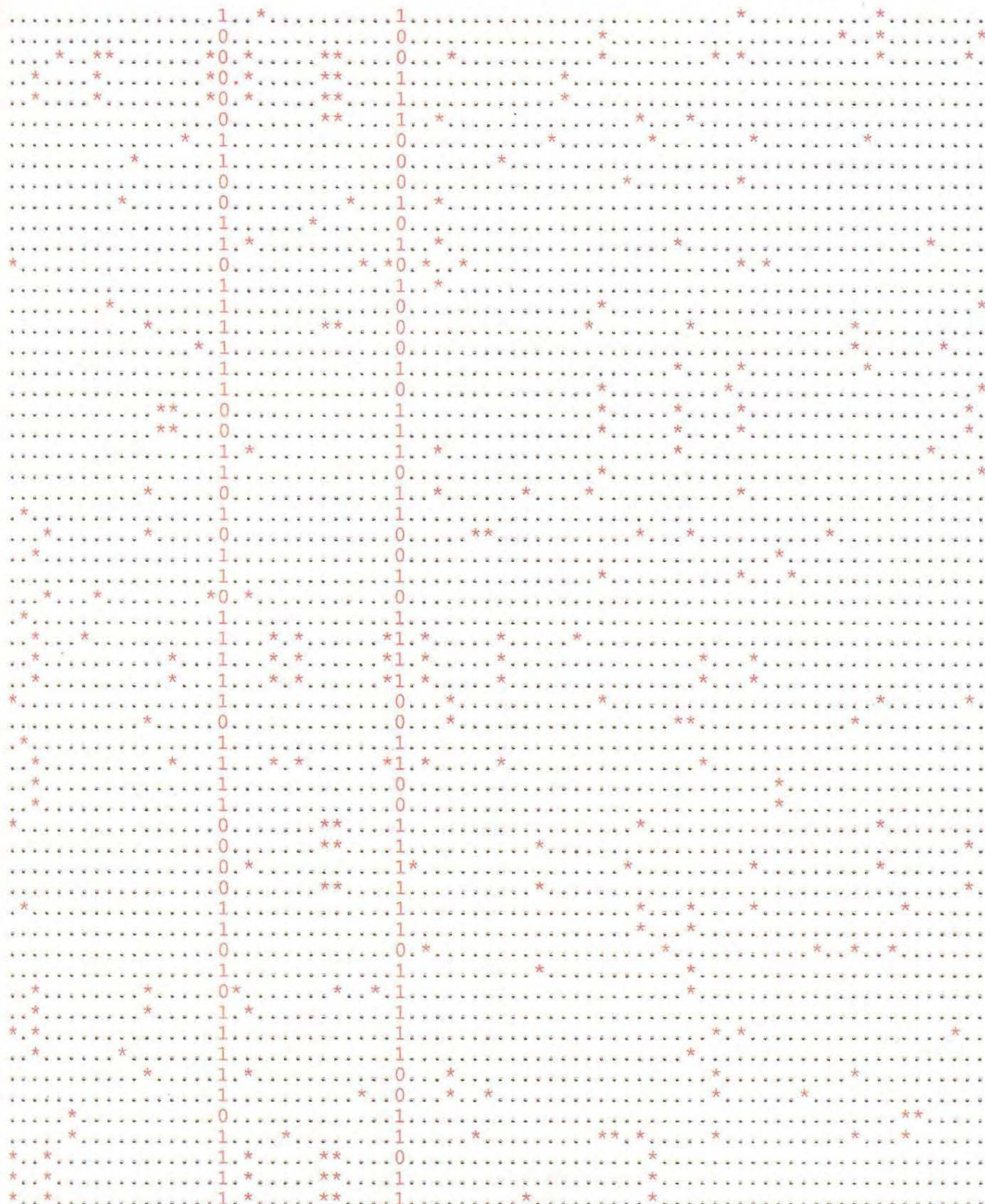


FIG. 17.4 O resultado de uma cortadora de quatro de 58 cromossomos, sondado quanto ao gene de xantina desidrogenase em *Drosophila pseudoobscura*. Cada linha é um cromossomo (haplótipo) amostrado de uma população natural. Cada posição ao longo da linha é uma taxa de restrição polimórfica ao longo da seqüência de 4,5 kb estudada. Onde aparece um asterisco, o haplótipo difere da maioria, seja cortando onde a maioria dos haplótipos não são cortados, ou não cortando onde a maioria dos haplótipos são cortados. Em dois sítios, não há um claro tipo majoritário, e, assim, 0 ou 1 é usado para mostrar se o sítio está ausente ou presente.

restrição ao longo de uma seqüência de 4,5 kb de tamanho. Marcantemente, entre 58 padrões, existem 53 diferentes. (Tente encontrar os pares idênticos.)

Repetições em tandem Os levantamentos dos fragmentos de restrição podem revelar uma outra forma de variação de seqüência de DNA, que surge da ocorrência de seqüências de DNA multiplamente repetidas. No genoma humano, uma variedade de seqüências curtas de DNA são dispersas pelo genoma, cada uma das quais é multiplamente repetida em seqüência (em *tandem*). O número de repetições pode variar de uma dúzia a mais de 100 em genomas individuais diferentes. Tais seqüências são conhecidas como **número variável de repetições em tandem (VNTR)**. Se as enzimas de restrição cortam seqüências flangeadoras de ambos os lados de tal arranjo em tandem, será produzido um fragmento cujo tamanho é proporcional ao número de elementos repetidos. Os fragmentos de tamanhos diferentes irão migrar em

velocidades diferentes em um gel de eletroforese. As cópias individuais dos elementos de seqüências repetidas são muito pequenas para distinguir entre, digamos, 64 e 68 repetições, mas os tamanhos das classes que incluem vários números repetidos (*bins*) podem ser estabelecidos, e uma população pode ser triada quanto às freqüências de classes diferentes. O Quadro 17.5 mostra os dados de duas VNTR amostradas em dois grupos de índios do Brasil. Em um caso, D14S1, os karitianas são quase homozigotos, enquanto os suruí são muito variáveis; no outro caso, D14S13, ambas as populações são variáveis, mas com padrões de freqüências diferentes.

Variação de seqüência completa Uma forma ubíqua de variação genética é a variação no nucleotídeo em uma única posição, chamada de **polimorfismo de nucleotídeo único (SNP)**. Os estudos de variação a nível de pares de bases únicos por seqüenciamento de DNA podem dar informações de dois tipos. Primeiro, a variação de seqüência do DNA pode

Quadro 17.5 Tamanho das Frequências de Classes para Duas Sequências Diferentes de VNTR, D14S1 e D14S13, em Karitianas e Suruís do Brasil

Tamanho da classe	D14S1		D14S13	
	Karitianas	Suruís	Karitianas	Suruís
3-4	105	41	0	0
4-5	0	3	3	14
5-6	0	11	1	4
6-7	0	2	1	2
7-8	0	1	1	2
8-9	3	3	8	16
9-10	0	11	28	9
10-11	0	2	22	0
11-12	0	4	18	8
12-13	0	0	13	18
13-14	0	0	13	3
>14	0	0	0	2
	108	78	108	78

Fonte: Dados de J. Kidd and K. Kidd, *Am. J. Phy. Anthro.* 81, 1992, 249.

ser estudada nas regiões codificantes de proteínas dos genes. As sequências de regiões codificantes podem ser traduzidas para revelar a sequência exata de diferenças de aminoácidos nas proteínas de indivíduos diferentes em uma população ou de espécies diferentes. O sequenciamento do DNA é superior em precisão aos estudos eletroforéticos de uma proteína de indivíduos diferentes, que podem mostrar que há variação nas sequências de aminoácidos, mas não podem identificar quantos ou quais aminoácidos diferem entre indivíduos. Assim, ao serem obtidas sequências de DNA de muitas variantes eletroforéticas de esterase-5 em *D. pseudoobscura* (veja a Figura 17.1), foram encontradas variantes eletroforéticas para diferenciar umas das outras por uma média de 8 aminoácidos, e os 20 tipos diferentes de aminoácidos estavam envolvidos em polimorfismos com a frequência aproximada na qual eles estavam presentes na proteína. Tais estudos também mostram que regiões diferentes da mesma proteína têm quantidades diferentes de polimorfismo. Para a proteína esterase-5, que consiste em 545 aminoácidos, 7% das posições de aminoácidos são polimórficas, mas os últimos aminoácidos no terminal carboxila da proteína são totalmente invariantes entre moscas individuais, provavelmente porque esses aminoácidos são necessários para o funcionamento apropriado da proteína.

Segundo, a variação da sequência de DNA também pode ser estudada nesses pares de bases que *não* determinam nem mudam a sequência da proteína. Tal variação de par de bases pode ser encontrada no DNA em 5' flanqueando sequências que podem ser regulatórias. A importância do estudo da variação em sequências regulatórias não pode ser supervalorizada. Foi sugerido que a maioria da evolução da forma, fisiologia e comportamento é baseada em mudanças de sequências reguladoras. Se isso é verdade, então grande

parte da variação de sequência em regiões codificantes e nas sequências de aminoácidos que elas codificam está além do ponto. Existe também variação em íntrons, em DNA não-transcrito a 3' do gene e nas posições dos nucleotídeos nos códons (geralmente terceiras posições) cuja variação não resulta em substituições de aminoácidos. Esses chamados polimorfismos de pares de bases *silenciosos* ou *sinônimos* são muito mais comuns do que as mudanças que resultam em polimorfismo de aminoácidos, supostamente porque muitas mudanças de aminoácidos interferem na função normal da proteína e são eliminadas por seleção natural.

Um exame do quadro de tradução de códons (veja a Figura 9.6) mostra que aproximadamente 25% de todas as mudanças aleatórias de pares de bases seriam sinônimas, dando um códon alternativo para o mesmo aminoácido, enquanto 75% das mudanças aleatórias mudariam o aminoácido codificado. Por exemplo, uma mudança de AAT para AAC ainda codifica asparagina, mas uma mudança para ATT, ACT, AAA, AAG, AGT, TAT, CAT ou GAT, todas mudanças de um único par de bases AAT, muda o aminoácido codificado. Assim, se as mutações de pares de bases fossem aleatórias e se a substituição de um aminoácido não fizesse diferença na função, esperaríamos uma proporção de 3:1 de substituição de aminoácidos para polimorfismos silenciosos. As proporções reais encontradas em *Drosophila* variam de 2:1 a 1:10. Claramente, há um grande excesso de polimorfismos sinônimos, mostrando que a maioria das mudanças de aminoácidos fazem uma diferença em função e, portanto, estão sujeitas a seleção natural. Não devemos supor, entretanto, que mutações silenciosas em sequências codificantes são inteiramente livres de restrições. Trincas alternativas diferentes codificantes do mesmo aminoácido podem diferir em velocidade e precisão de transcrição, e o mRNA correspondente a diferentes trincas alternativas

podem ter uma precisão e velocidade de tradução diferentes devido a limitações do conjunto de tRNA disponíveis. A evidência deste último efeito é que trincas sinônimas alternativas para um aminoácido não são igualmente usadas, e a desigualdade de uso é muito mais pronunciada para genes que são transcritos a uma velocidade muito alta.

No DNA de um gene, existem também restrições nas seqüências não-codificantes 5' e 3' e nas seqüências íntron. Tanto as seqüências de DNA 5' e 3' não-codificantes contêm sinais para transcrição, e íntrons podem conter acentuadores de transcrição (veja o Capítulo 11).

Dentro das espécies, há uma grande variação genética. Essa variação se manifesta como a forma e o número dos cromossomos, e como segmentos de DNA que não têm efeitos desenvolvimentais observáveis.

► 17.2 Efeitos da Reprodução Sexual na Variação

Segregação meiótica e equilíbrio genético

Se a herança fosse baseada em uma substância contínua como o sangue, então a reprodução de indivíduos com fenótipos diferentes produziria uma prole que seria de fenótipo intermediário. Quando esses tipos intermediários fossem cruzados, sua prole novamente seria intermediária. Uma população na qual os indivíduos se reproduzem aleatoriamente perde lentamente toda a sua variação, e cada membro da população teria o mesmo genótipo.

A natureza particulada da herança muda completamente esse quadro. Devido à natureza distinta dos genes e à segregação de alelos na meiose, um cruzamento entre indivíduos intermediários *não* resulta em toda a prole intermediária. Ao contrário, parte da prole será de tipos extremos, os que são homozigotos. Considere uma população na qual machos e fêmeas se reproduzem aleatoriamente com relação ao locus gênico *A*; isto é, o genótipo nesse locus não é um fator na escolha do parceiro. Tal reprodução aleatória é equivalente a misturar todos os espermatozoides e ovócitos na população e, então, juntar aleatoriamente cada espermatozoide com um ovócito colhido de modo aleatório.

O resultado de tal pareamento aleatório de espermatozoides e ovócitos é fácil de calcular. Se, em alguma população, a frequência do alelo *A* é 0,60 tanto nos espermatozoides quanto nos ovócitos, então a chance de um espermatozoide escolhido aleatoriamente e um ovócito colhido aleatoriamente serem ambos *A* é $0,60 \times 0,60 = 0,36$. Assim, em uma população de reprodução aleatória com essa frequência alélica, 36% da prole seriam *A/A*. Do mesmo modo, a frequência de prole *a/a* seria de $0,40 \times 0,40 = 0,16$. Os heterozigotos seriam produzidos pela fusão ou de um espermatozoide *A* com um ovócito *a*, ou de um espermatozoide *a* com um ovócito *A*. Se os gametas formam pares aleatoriamente, então a chance de um espermatozoide *A* e um ovócito *a* é de $0,60 \times 0,40$, e a combinação reversa tem a mesma

probabilidade; logo, a frequência de prole heterozigota é $2 \times 0,60 \times 0,40 = 0,48$.

Agora podemos compreender por que a variação é mantida em uma população. O processo de pareamento aleatório não fez nada para mudar as frequências *alélicas*, como pode ser facilmente verificado calculando as frequências dos alelos *A* e *a* entre a prole nesse exemplo, usando o método descrito no Boxe 17.1. Assim, as proporções de homozigotos e heterozigotos em cada geração sucessiva permanecerão as mesmas. Essas frequências constantes formam a **distribuição de equilíbrio**. O Boxe 17.2 dá a forma geral desse resultado em equilíbrio.

A segregação meiótica em populações que se reproduzem aleatoriamente resulta em uma distribuição de equilíbrio de genótipos após apenas uma geração, e assim é mantida a variação genética.

A distribuição de equilíbrio pode ser calculada de acordo com a fórmula

$$\begin{array}{ccc} A/A & A/a & a/a \\ p^2 & 2pq & q^2 \end{array}$$

quando *p* é a frequência do alelo *A*, *q* é a frequência do alelo *a*, e $p + q = 1$.

Essa distribuição é chamada de **equilíbrio de Hardy-Weinberg** em homenagem a Godfrey Hardy e Wilhelm Weinberg, as duas pessoas que, independentemente, descobriram esses resultados de reprodução aleatória em um equilíbrio de frequências de genótipo em uma população. (Uma terceira descoberta independente foi feita pelo geneticista russo Sergei Chetverikov.)

O equilíbrio de Hardy-Weinberg significa que a reprodução sexual não causa uma redução constante na variação genética de cada geração; ao contrário, a quantidade de variação permanece constante geração após geração, na ausência de outras forças perturbadoras. O equilíbrio é a consequência direta da segregação de alelos na meiose em heterozigotos.

Numericamente, o equilíbrio mostra que, independentemente da mistura particular de genótipos na geração parental, a distribuição genotípica, após uma rodada de reprodução aleatória, é completamente especificada pela frequência alélica *p*. Por exemplo, considere três populações hipotéticas que surgiram da mistura de migrantes de fontes diferentes:

	$f(A/A)$	$f(A/a)$	$f(a/a)$
I	0,3	0,0	0,7
II	0,2	0,2	0,6
III	0,1	0,4	0,5

A frequência alélica *p* de *A* em todas as três populações é

$$\begin{array}{ll} \text{I} & p = f(A/A) + \frac{1}{2} f(A/a) \\ & = 0,3 + \frac{1}{2} (0) = 0,3 \\ \text{II} & p = 0,2 + \frac{1}{2} (0,2) = 0,3 \\ \text{III} & p = 0,1 + \frac{1}{2} (0,4) = 0,3 \end{array}$$

Boxe 17.2 O Equilíbrio de Hardy-Weinberg

Se a frequência do alelo A é p tanto nos espermatozóides quanto nos ovócitos, a frequência do alelo a é $q = 1 - p$, e, então, as consequências das uniões aleatórias de espermatozóides e ovócitos são as mostradas no diagrama seguinte. A probabilidade de que tanto os espermatozóides quanto os ovócitos na reprodução levem A é

$$p \times p = p^2$$

e assim será a frequência de homozigotos A/A na geração seguinte. Similarmente, a chance de heterozigotos A/a será

$$(p \times q) + (q \times p) = 2pq$$

e a chance dos homozigotos a/a será

$$q \times q = q^2$$

Os três genótipos, após uma geração de reprodução aleatória, estarão nas frequências

$$p^2 : 2pq : q^2$$

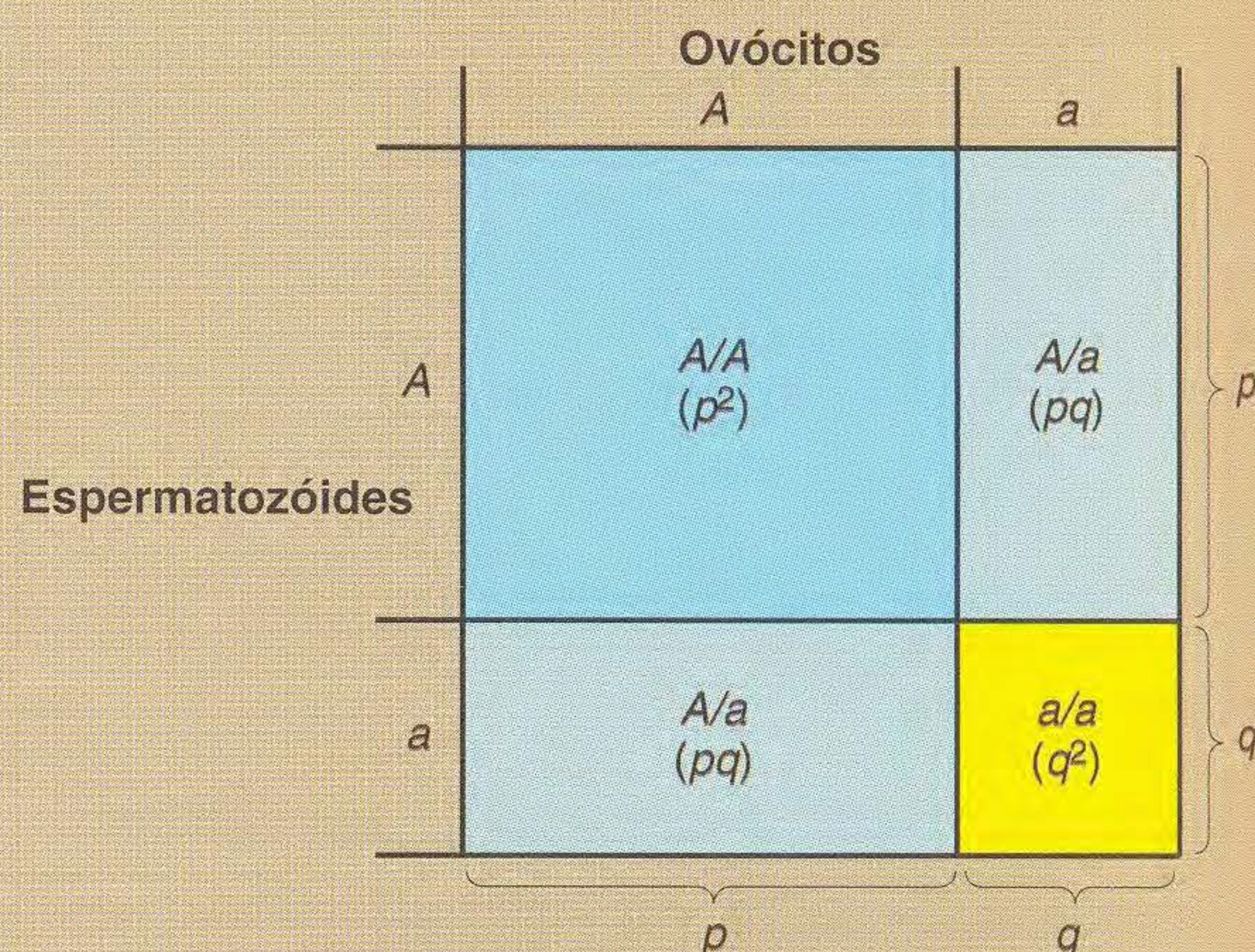
A frequência de A na F_1 não mudará (ela ainda será p), pois, como mostra o diagrama, a frequência de A nos zóigotos é a frequência de A/A mais metade da frequência de A/a , ou

$$p^2 + pq = p(p + q) = p$$

Portanto, na segunda geração, as frequências dos três genótipos serão novamente

$$p^2 : 2pq : q^2$$

e assim por diante, para sempre. Essas são as frequências de equilíbrio de Hardy-Weinberg.



As frequências do equilíbrio de Hardy-Weinberg que resultam de reprodução aleatória.

Assim, a despeito de suas composições genóticas muito diferentes, elas têm a mesma frequência alélica. Após uma geração de reprodução aleatória dentro de cada população, entretanto, cada uma delas terá as mesmas frequências genóticas:

$$A/A: p^2 = (0,3)^2 = 0,09$$

$$A/a: 2pq = 2(0,3)(0,7) = 0,42$$

$$a/a: q^2 = (0,7)^2 = 0,49$$

e permanecerão assim indefinidamente.

Uma consequência das proporções de Hardy-Weinberg é que alelos raros quase nunca estão em condição homozigota. Um alelo com uma frequência de 0,001 está presente em homozigose com uma frequência apenas de 1 em um milhão; a maioria das cópias de tais alelos raros é encontrada em heterozigotos. Podemos calcular a frequência relativa do alelo em heterozigotos (em contraste com homozigotos) usando as frequências do equilíbrio de Hardy-Weinberg. Em geral, duas cópias de um alelo estão presentes em homozigotos, mas apenas uma cópia desse alelo está presente em cada heterozigoto; logo, a frequência relativa do alelo q em heterozigotos é dada pela proporção da frequência de heterozigotos $2pq$ para o dobro da frequência do homozigoto q^2 :

$$\frac{2pq}{2q^2} = \frac{p}{q}$$

que para $q = 0,001$ é uma proporção de 999:1. Por exemplo, uma mutação recessiva causa a doença potencialmente letal

fenilcetonúria (PKU) quando em homozigose. Assim, o alelo mutante tem uma frequência de 0,01 em populações européias e americanas. A frequência da doença, entretanto, é apenas de cerca de 1 em 10.000 nascimentos. A relação geral entre as frequências de homozigotos e heterozigotos como uma função das frequências alélicas é mostrada na Figura 17.5.

Em nossa derivação do equilíbrio, supomos que a frequência alélica p é a mesma nos espermatozóides e nos ovócitos. O teorema do equilíbrio de Hardy-Weinberg não se aplica a genes ligados ao sexo se machos e fêmeas começarem com frequências desiguais.

O princípio do equilíbrio de Hardy-Weinberg pode ser generalizado para incluir casos onde há mais de dois alelos na população. Em geral, não importa quantos tipos alélicos estejam presentes na população, a frequência dos homozigotos de determinado alelo é igual ao quadrado da frequência do alelo. A frequência de heterozigotos para um par particular de alelos é o dobro do produto da frequência desses dois alelos. Por exemplo, suponha que existam três alelos, A_1 , A_2 , e A_3 , cujas frequências são 0,5, 0,3 e 0,2, respectivamente. Então, as frequências do equilíbrio de Hardy-Weinberg dos homozigotos serão

A_1A_1	A_2A_2	A_3A_3
$(0,5)^2 = 0,25$	$(0,3)^2 = 0,09$	$(0,2)^2 = 0,04$

As frequências homozigotas e heterozigotas dependem de frequências alélicas

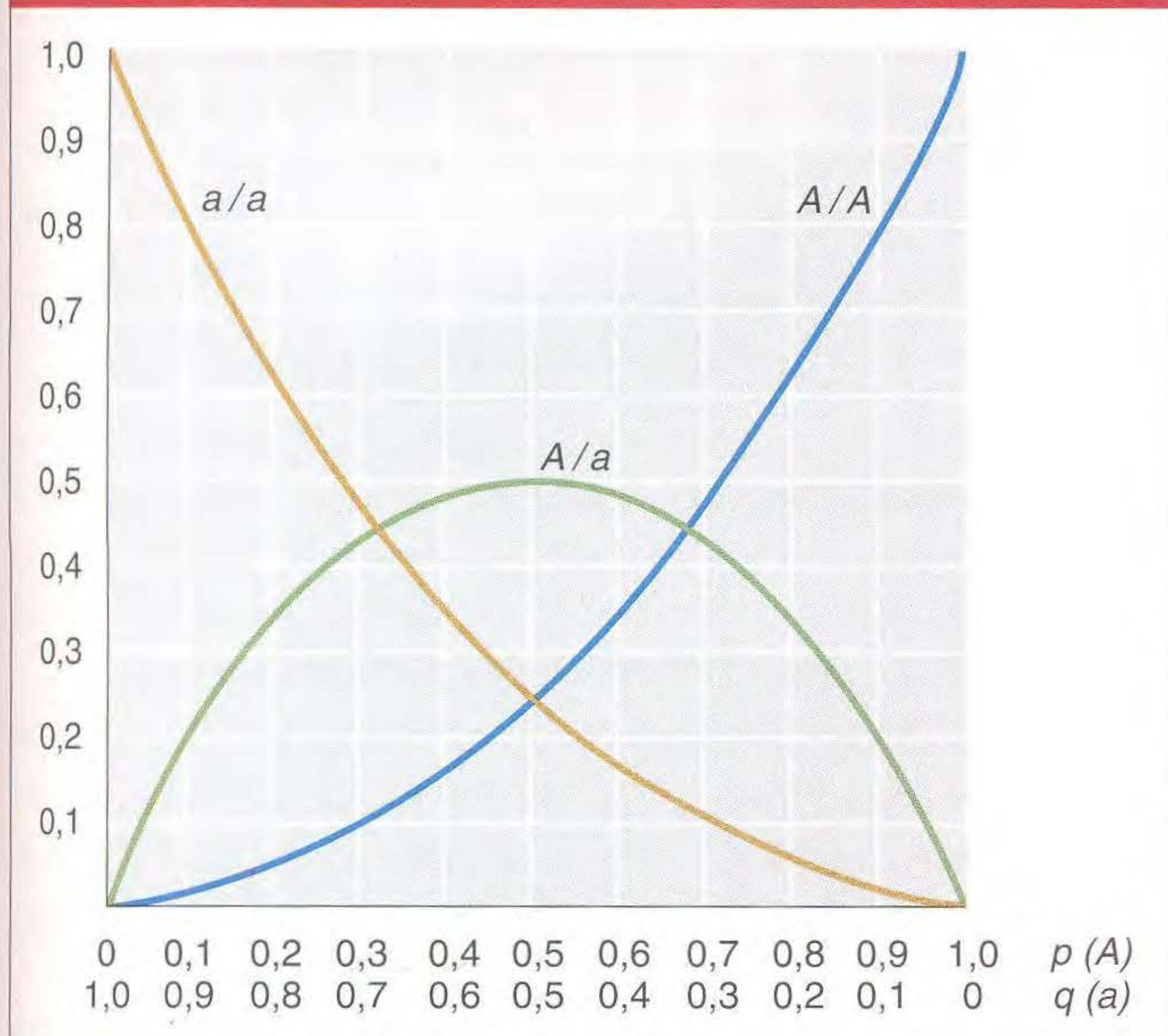


FIG. 17.5 Frequências homozigotas e heterozigotas em função das frequências alélicas. As curvas mostrando as proporções dos homozigotos A/A (linha azul), homozigotos a/a (linha laranja) e heterozigotos A/a (linha verde) em populações com frequências alélicas diferentes se as populações estiverem no equilíbrio de Hardy-Weinberg.

e as frequências de heterozigotos serão

$$\begin{array}{ccc} A_1A_2 & A_1A_3 & A_2A_3 \\ 2(0,5)(0,3) = 0,30 & 2(0,5)(0,2) = 0,20 & 2(0,3)(0,2) = 0,12 \end{array}$$

Heterozigose

Uma medida da variação genética (em contraste com sua *descrição* pelas frequências alélicas) é a quantidade de **heterozigose** para um gene em uma população, que é dada pela frequência total de heterozigotos do gene. Essa heterozigose pode ser ou observada diretamente, contando os heterozigotos, ou calculada pelas frequências alélicas, usando as proporções do equilíbrio de Hardy-Weinberg. Se um alelo está com uma frequência muito alta e todos os outros são quase zero, então haverá pouca heterozigose porque a maioria dos indivíduos serão homozigotos para o alelo comum. Esperamos

que a heterozigose seja maior quando existem muitos alelos de um gene, todos em igual frequência. No Quadro 17.1, a heterozigose é simplesmente igual à frequência do genótipo M/N em cada população.

Quando é considerado mais de um locus, existem dois modos possíveis para calcular a heterozigose. O gene S (que codifica o fator secretor, determinando se as proteínas M e N também estão contidas na saliva) está proximamente ligado a M/N em humanos. O Quadro 17.6 mostra as frequências de quatro combinações dos dois alelos para os dois genes ($M S$, $M s$, $N S$ e $N s$) em várias populações. Para o primeiro modo de medir a heterozigose, podemos calcular a frequência de heterozigotos em cada locus separadamente (heterozigose alélica). Para o segundo modo, podemos considerar se os cromossomos homólogos de um indivíduo levam à mesma combinação de alelos. A combinação de alelos de genes diferentes no mesmo cromossomo homólogo é chamada de **haplótipo**. Para determinar se os dois alelos de genes diferentes estão associados no mesmo cromossomo homólogo, é necessário sequenciar o DNA de indivíduos ou ter informações sobre seus genitores ou prole. Quando temos essa informação, consideramos cada haplótipo como uma unidade, como no Quadro 17.6, e calculamos a proporção de todos os indivíduos que levam duas formas haplotípicas diferentes. Essa forma de heterozigose também é chamada de *diversidade de haplótipo*. Os resultados de ambos os cálculos são dados no Quadro 17.6. Note que a diversidade de haplótipos é sempre maior do que a heterozigose média de loci separados, porque um indivíduo é heterozigoto haplotípico se *ambos* os seus genes forem heterozigotos.

Reprodução aleatória

O equilíbrio de Hardy-Weinberg foi obtido com suposições de “reprodução aleatória”, mas devemos distinguir cuidadosamente dois significados desse processo. Primeiro, podemos significar que os indivíduos não escolhem seus parceiros com base em alguma característica herdável. Os seres humanos reproduzem-se aleatoriamente com relação a grupos sanguíneos nesse sentido porque, em geral, não sabem os tipos de seus prospectivos pares, e, mesmo que saibam, é improvável que usem o tipo sanguíneo como critério de escolha. Nesse primeiro sentido, a reprodução aleatória ocorrerá com relação a genes que não têm efeito no aspecto, comportamento, odor, ou outras características que influenciam diretamente a escolha do parceiro.

Quadro 17.6 Frequências de Tipos de Haplótipos para Sistema MNS em Várias Populações Humanas

População	Haplótipo				Heterozigose (H)	
	$M S$	$M s$	$N S$	$N s$	De haplótipos	De alelos
Ainu	0,024	0,381	0,247	0,348	0,672	0,438
Ugandenses	0,134	0,357	0,071	0,438	0,658	0,412
Paquistaneses	0,177	0,405	0,127	0,291	0,704	0,455
Ingleses	0,247	0,283	0,080	0,290	0,700	0,469
Navaho	0,185	0,702	0,062	0,051	0,467	0,286

Fonte: A.E. Mourant, *The Distribution of the Human Blood Groups*. Blackwell Scientific, 1954.

Em segundo lugar, a reprodução aleatória é relevante quando há uma divisão de uma espécie em subgrupos. Suponha que as frequências de alelos diferem de grupo para grupo. Se os indivíduos tendem a se reproduzir dentro de seu próprio grupo (**endogamia**), então a reprodução não é aleatória com relação à espécie como um todo, e as frequências de genótipos se afastarão mais ou menos das frequências de Hardy-Weinberg. Nesse sentido, os seres humanos não estão se reproduzindo aleatoriamente, pois grupos étnicos e raciais e populações separadas geograficamente diferem umas das outras em frequências gênicas, e as pessoas apresentam altas taxas de endogamia não só dentro de importantes raças geográficas, mas também dentro de grupos étnicos locais. Os espanhóis e os russos diferem em suas frequências de grupo sanguíneo ABO. Os espanhóis geralmente se casam com espanhóis, e os russos com russos, e há uma reprodução não-aleatória não-intencional com relação a grupos sanguíneos ABO.

O Quadro 17.7 apresenta reprodução aleatória nesse primeiro sentido e não-aleatória no segundo sentido do grupo sanguíneo MN. Nas subpopulações de esquimós, egípcios, chineses e australianos, as mulheres não escolhem seus parceiros pelo tipo MN, e, portanto, o equilíbrio de Hardy-Weinberg existe *dentro* de subpopulações. Mas os egípcios em geral não se reproduzem com esquimós, ou aborígenes australianos, e, assim, as associações não-aleatórias na espécie humana *como um todo* resultam em grandes diferenças nas frequências genotípicas de grupo para grupo. Ocorre que, se considerarmos a espécie humana como um todo e pudermos calcular a frequência alélica média em toda a espécie, observaremos um desvio do equilíbrio de Hardy-Weinberg para a espécie. Entretanto, para fazermos tal cálculo, precisamos conhecer o tamanho da população e as frequências alélicas de cada população local. Para ilustrar o efeito, suponha que juntemos um grupo de números iguais de esquimós e aborígenes australianos. Pelas frequências genotípicas observadas no Quadro 17.7, podemos calcular que as frequências alélicas nos dois subgrupos e o grupo misto são

Se o grupo misto fosse realmente uma única população com reprodução aleatória, esperaríamos encontrar as proporções de Hardy-Weinberg dadas pelas frequências alélicas médias,

p^2 (M/M)	$2pq$ (M/N)	q^2 (N/N)
0,297	0,496	0,207

enquanto o que de fato encontramos é a proporção média de homozigotos e heterozigotos das duas populações parentais originais:

(M/M)	(M/N)	(N/N)
0,430	0,230	0,340

Endogamia e reprodução preferencial

A reprodução aleatória com relação a um locus é comum dentro das populações, mas não é universal. Devem ser diferenciados dois tipos de reprodução aleatória. Primeiro, os indivíduos podem se reproduzir com outros com os quais compartilham algum grau de ancestralidade comum, isto é, algum grau de correlação genética. Se a reprodução entre parentes é mais comum do que ocorreria por puro acaso, então a população é **endogâmica**. Se a reprodução entre parentes é menos comum do que ocorreria por acaso, então diz-se que a população está sofrendo **exogamia forçada** ou **endogamia negativa**.

Segundo, os indivíduos podem tender a escolher seus parceiros não porque são aparentados, mas devido à sua semelhança em alguma característica. A tendência de reprodução entre similares é chamada de **reprodução preferencial positiva**. O casamento com parceiros diferentes é chamado de **reprodução preferencial negativa**. O casamento preferencial nunca é completo; logo, em alguma população, alguns casamentos serão aleatórios e alguns resultarão de casamentos preferenciais.

Endogamia e casamento preferencial não são idênticos. Parentes próximos se assemelham muito mais do que indivíduos não-relacionados em média, mas não necessariamente por qualquer característica fenotípica particular em determinados indivíduos. Assim, a endogamia pode resultar na reprodução de indivíduos bem diferentes. Por outro lado, os indivíduos que se assemelham por alguma característi-

Quadro 17.7 Comparação Entre as Frequências Observadas de Genótipos para o Locus do Grupo Sanguíneo MN e as Frequências Esperadas de Reproduções Aleatórias

População	Observada			Esperada		
	M/M	M/N	N/N	M/M	M/N	N/N
Esquimós	0,835	0,156	0,009	0,834	0,159	0,008
Egípcios	0,278	0,489	0,233	0,274	0,499	0,228
Chineses	0,332	0,486	0,182	0,331	0,488	0,181
Aborígenes australianos	0,024	0,304	0,672	0,031	0,290	0,679

Nota: As frequências esperadas são computadas de acordo com o equilíbrio de Hardy-Weinberg, usando os valores de p e q computados das frequências observadas.

ca podem fazê-lo porque são parentes, mas indivíduos não-aparentados também podem ter semelhanças específicas. Irmãos e irmãs não têm sempre a mesma cor de olhos, e as pessoas de olhos azuis não são todas aparentadas.

A reprodução preferencial para algumas características é comum. Em humanos, há uma tendência de casamento preferencial para cor da pele e altura, por exemplo. Uma diferença importante entre casamento preferencial e endogamia é que o primeiro é específico a um determinado fenótipo, enquanto o segundo aplica-se a todo o genoma. Os indivíduos podem se reproduzir preferencialmente com relação à altura, mas aleatoriamente quanto ao grupo sanguíneo. Os primos em primeiro grau, por outro lado, assemelham-se geneticamente na média no mesmo grau em todos os loci.

Para ambos, casamento preferencial e endogamia, a consequência para a estrutura da população é a mesma: há um aumento de homozigose acima do nível previsto pelo equilíbrio de Hardy-Weinberg. Se dois indivíduos são aparentados, eles têm pelo menos um ancestral comum. Assim, há alguma chance de um alelo levado por um deles e um alelo levado pelo outro serem ambos descendentes de uma molécula idêntica de DNA. O resultado é que há uma chance extra de essa **homozigose por descendência** ser somada à chance de homozigose ($p^2 + q^2$) que surge por reprodução aleatória de indivíduos não-aparentados. A probabilidade dessa homozigose extra por descendência é chamada de **coeficiente de endogamia**, F . A Figura 17.6 ilustra o cálculo da probabilidade por descendência. Os indivíduos I e II são irmãos porque compartilham ambos os genitores. Marcamos cada alelo nos genitores diferentemente para podermos rastreá-los. Os indivíduos I e II reproduzem-se para produzir III. Suponha

que o indivíduo I seja A_1/A_3 e o gameta que ele contribui para III contenha o alelo A_1 ; então gostaríamos de calcular a probabilidade de que o gameta produzido por II também seja A_1 . A chance é $1/2$ de que II receba A_1 de seu pai, e, se isso ocorrer, a chance é $1/2$ de que II passe A_1 para o gameta em questão. Assim, a probabilidade de que III receba um A_1 de II é $1/2 \times 1/2 = 1/4$, que é a chance de que III, o produto da reprodução de irmãos, seja homozigoto A_1/A_1 por descendência do ancestral original.

Tal endogamia próxima pode ter consequências deletérias. Consideremos um raro alelo deletério a que, quando homozigoto, causa um distúrbio metabólico. Se a frequência do alelo na população é p , então a probabilidade de que um casal aleatório produza uma prole homozigota é apenas de p^2 (do equilíbrio de Hardy-Weinberg). Assim, se p é, digamos $1/1.000$, a frequência de homozigotos será 1 em 1.000.000. Agora, suponha que o casal seja irmão e irmã. Se um dos genitores em comum é um heterozigoto para a doença, ambos podem ter recebido o alelo deletério e passá-lo para sua prole. Como o cálculo mostra, há uma chance de $1/4$ de que a prole de uma reprodução irmão-irmã seja homozigota para um dos alelos levados por seus avós. Suponha que, entre as quatro cópias do gene possuído pelos avós, uma era uma mutação deletéria. Portanto, a chance de que uma prole da reprodução irmão-irmã seja homozigota para o alelo deletério é $1/4 \times 1/4 = 1/16$. Existem tantos alelos raros deletérios para genes diferentes nas populações humanas que cada um de nós é heterozigoto para muitos desses alelos raros. Assim, são muito altas as chances de que uma prole de reprodução irmão-irmã seja homozigota para pelo menos um deles.

A endogamia sistemática entre parentes próximos acaba levando à homozigose completa da população, mas em taxas diferentes, dependendo do grau de parentesco. Se dois alelos estão presentes em uma população endogâmica, um eventualmente será perdido e o outro terá uma frequência de 1,0. Em outras palavras, ele se tornará **fixado**. Para alelos que não são deletérios, qual alelo será fixado é uma questão do acaso. Suponha, por exemplo, que vários grupos de indivíduos sejam colhidos de uma população e submetidos a endogamia. Se, na população original da qual foram colhidas as linhagens endogâmicas, o alelo A tem uma frequência p e o alelo a tem uma frequência $q = 1 - p$, então a proporção p das linhagens homozigotas estabelecidas por endogamia será homozigota A/A e uma proporção q das linhagens será a/a . A endogamia usa a variação genética presente *dentro* da população original e a converte em variação *entre* as linhagens endogâmicas homozigotas obtidas da população (Figura 17.7).

Consideremos como a endogamia leva a uma perda de variação. Suponha que uma população é fundada por um pequeno número de indivíduos que se reproduzem aleatoriamente para produzir a geração seguinte. Suponha que não ocorreu mais nenhuma outra migração para a população. (Por exemplo, os coelhos hoje na Austrália provavelmente descenderam de uma única introdução de alguns animais no século XIX.) Muito embora a reprodução seja aleatória dentro da população, nas gerações posteriores todos são aparentados entre si, pois suas árvores genealógicas têm ancestrais comuns em alguns pontos

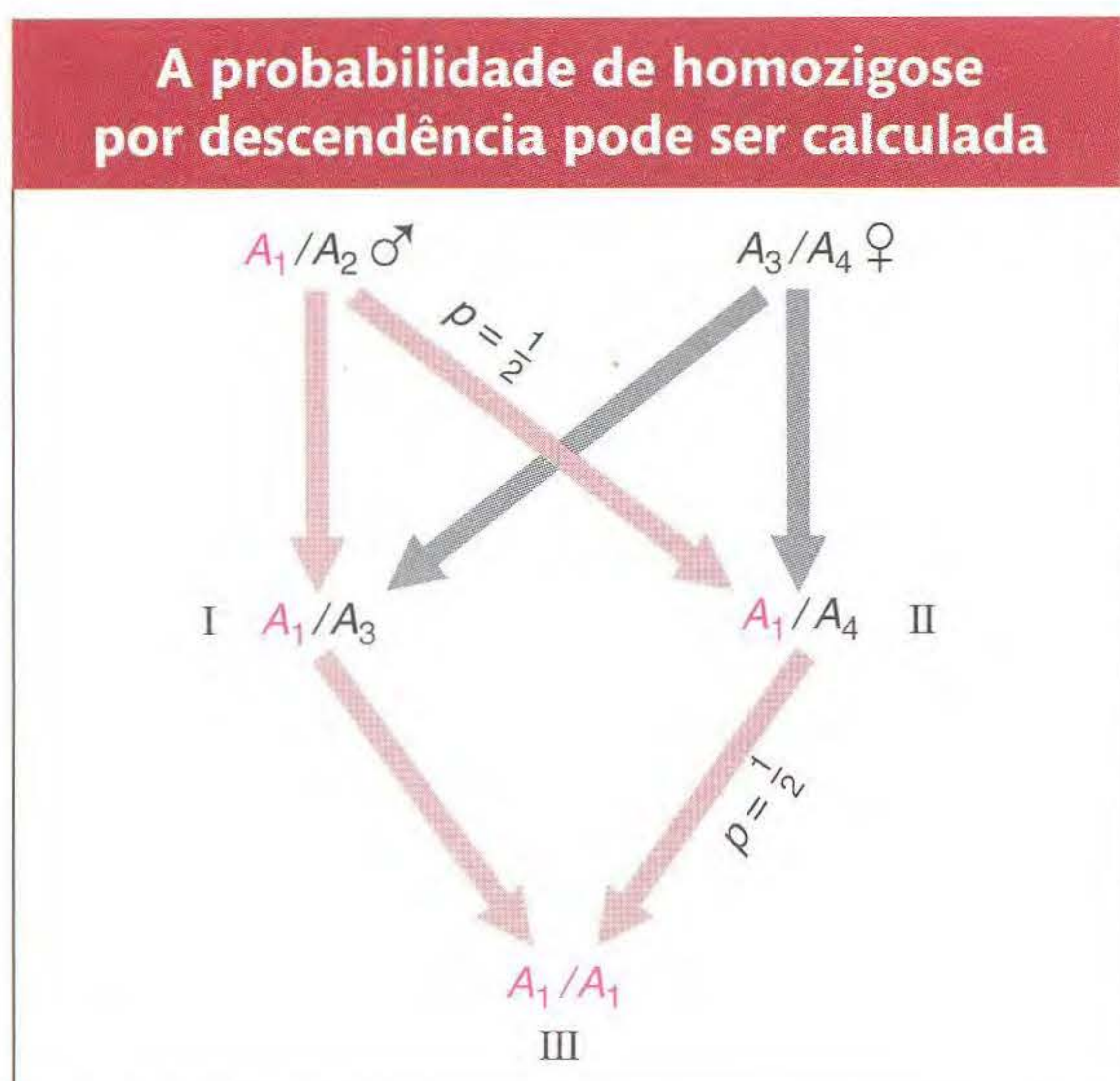


FIG. 17.6 Cálculo de homozigose por descendência para uma prole (III) de reprodução entre irmão-irmã (I-II). Suponha que o indivíduo III recebeu uma cópia de A_1 de seu avô através do indivíduo I. A probabilidade de que II receba A_1 de seu pai é $1/2$; se tiver, a probabilidade de que II passe A_1 para III é $1/2$. Assim, a probabilidade de que III receba um A_1 de II é $1/2 \times 1/2 = 1/4$.

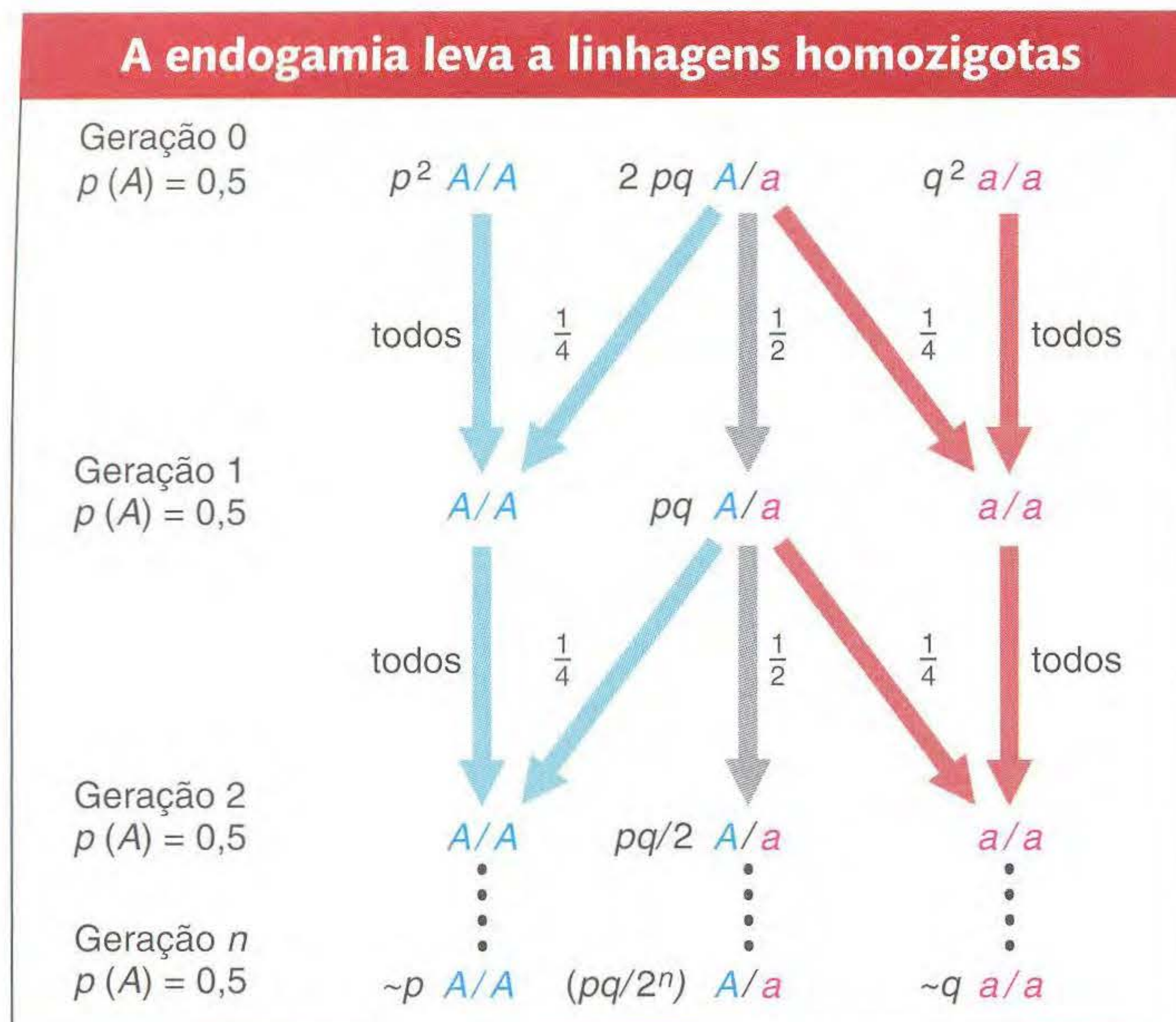


FIG. 17.7 Gerações repetidas de endogamia (ou autofecundação) acabarão dividindo uma população heterozigota em uma série de linhagens completamente homozigotas. A frequência de linhagens A/A entre as homozigotas será igual à frequência (p) do alelo A na população heterozigota original, enquanto a frequência de linhagens a/a será igual à frequência original de a (q).

em seus pedigrees. Tal população é então endogâmica no sentido de que há alguma probabilidade de um gene ser homozigoto por descendência. Como a população é, necessariamente, de tamanho finito, algumas das linhagens familiares originalmente introduzidas serão extintas a cada geração, do mesmo modo que os nomes de famílias desaparecem em uma população humana que nunca recebe imigrantes, pois, por acaso, não foi deixada uma prole masculina. À medida que as famílias originais desaparecem, a população torna-se composta de cada vez menos indivíduos fundadores originais, e todos os membros da população tornam-se cada vez mais portadores dos mesmos alelos por descendência. Em outras palavras, o coeficiente de endogamia F aumenta, e a heterozigose diminui com o passar do tempo, até que, finalmente, F atinja 1,00 e a heterozigose chegue a 0.

A taxa de perda de heterozigose por geração em tal população de reprodução aleatória, fechada, finita, é inversamente proporcional ao número ($2N$) de genomas haplóides, onde N é o número de indivíduos diplóides na população. A cada geração, $1/(2N)$ da heterozigose restante é perdida.

17.3 Fontes de Variação

Para determinada população, existem três fontes de variação: mutação, recombinação e imigração de genes. Entretanto, a recombinação entre genes por si só não produz variação, a menos que já exista variação alélica segregando em diferentes loci. De outro modo não haveria o que recombinar. Similarmente, a migração não pode dar variação se toda a espécie é homozigota para o mesmo alelo. Finalmente, a fonte de toda a variação deve ser a mutação.

Variação por mutação

As mutações são a *fonte* da variação, mas o *processo* de mutação em si não ativa a mudança genética nas populações. A taxa de mudança na frequência gênica do processo de mutação é muito baixa porque as taxas de mutação espontânea são baixas (Quadro 17.8). A taxa de mutação é definida como a probabilidade de que uma cópia de um alelo mude para alguma outra forma alélica em uma geração. Assim, o aumento na frequência de um alelo mutante será o produto da taxa de mutação vezes a frequência do alelo não-mutante. Suponha que uma população fosse completamente homozigota A e as mutações para a ocorressem a uma taxa de $1/100.000$ por gameta recém-formado. Então, na geração seguinte, a frequência de alelos a seria apenas $1,0 \times 1/100.000 = 0,00001$, e a frequência dos alelos A seria $0,99999$. Após uma outra geração de mutação, a frequência de a teria aumentado em $0,99999 \times 0,00001$ para uma nova frequência de $0,0000199999$, enquanto o alelo original teria sido reduzido em frequência para $0,9999800001$. A taxa de aumento do novo alelo é extremamente lenta, e *fica mais lenta a cada geração* porque existem poucas cópias do alelo antigo deixadas para mutar. Assim, quando a mutação atingisse uma frequência de, digamos, 10%, a mudança de frequência na geração seguinte como resultado de mutação seria apenas de $0,9 \times 0,00001 = 0,000009$, cerca de 0,1 do que era na primeira geração. Uma fórmula geral para a mudança na frequência alélica sob mutação é dada no Boxe 17.3.

As taxas de mutação são tão lentas que a mutação sozinha não pode responder pelas rápidas mudanças genéticas nas populações e espécies.

Quadro 17.8 Taxas de Mutação de Ponto em Organismos Diferentes

Organismo	Gene	Taxa de mutação por geração
Bacteriófago	Gama de hospedeiros	$2,5 \times 10^{-9}$
<i>Escherichia coli</i>	Resistência do fago	2×10^{-8}
<i>Zea mays</i> (milho)	R (fator de cor)	$2,9 \times 10^{-4}$
	Y (sementes amarelas)	2×10^{-6}
<i>Drosophila melanogaster</i>	Média de letais	$2,6 \times 10^{-5}$

Fonte: T. Dobzhansky, *Genetics and the Origin of Species*, 3rd ed., rev. Columbia University Press, 1951.

Boxe 17.3**O Efeito da Mutação na Frequência Alélica**

Seja μ a **taxa de mutação** do alelo A para algum outro alelo a (a probabilidade de que um gene A se torne a durante a replicação do DNA na meiose precedente). Se p_t é a frequência do alelo A na geração t , $q_t = 1 - p_t$ é a frequência do alelo a na geração t e não há outros casos de mudança de frequência gênica (sem seleção natural, por exemplo), então a mudança na frequência alélica em uma geração é

$$\Delta p = p_t - p_{t-1} = (p_{t-1} - \mu p_{t-1}) - p_{t-1} = -\mu p_{t-1}$$

onde p_{t-1} é a frequência na geração precedente. Isso nos diz que a frequência de A diminui (e a frequência de a aumenta) em uma quantidade que é proporcional à taxa de mutação μ e à proporção p de todos os genes que ainda estão disponíveis para mutar. Assim Δp fica menor à medida que a frequência de p diminui, porque existem menos e menos alelos A para mutar em alelos a . Podemos fazer uma aproximação de que, após n gerações de mutação,

$$p_n = p_0 e^{-n\mu}$$

onde e é a base dos logaritmos naturais.

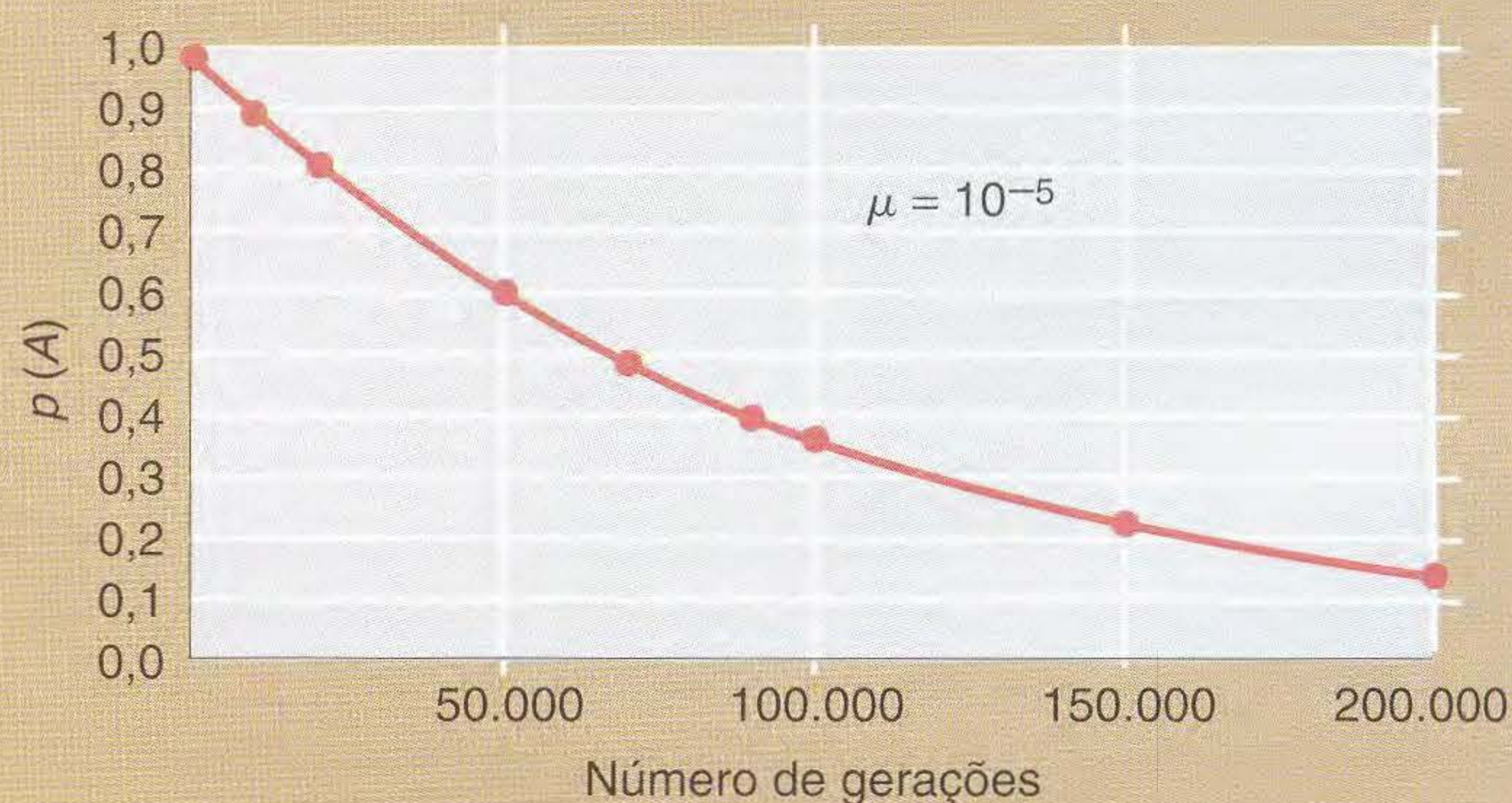
Essa relação de frequência alélica com o número de gerações é mostrada na figura anexa para $\mu = 10^{-5}$. Após 10.000 gerações de mutação contínua de A para a ,

$$p = p_0 \exp(10^4 \times 10^{-5}) = p_0 e^{-0,1} = 0,904 p_0$$

Se a população começa apenas com alelos A ($p_0 = 1,0$), ela ainda terá apenas 10% de alelos a após 10.000 gerações a

essa taxa de mutação relativamente alta e precisaria de 60.000 gerações adicionais para reduzir p para 0,5.

Mesmo que as taxas de mutação fossem duplicadas (digamos, por mutágenos ambientais), a taxa de mudança seria muito lenta. Por exemplo, os níveis de radiação de intensidade diferente para duplicar a taxa de mutação durante a vida reprodutiva de um ser humano são o limite das regulações de segurança ocupacional, e uma dose de radiação suficiente para aumentar as taxas de mutação em uma ordem de grandeza seria letal. Assim, a rápida mudança genética na espécie não seria um dos efeitos da radiação aumentada. Embora tenhamos muitas coisas a temer pela poluição da radiação ambiental, virarmos uma espécie de monstros não é uma delas.



A mudança em gerações na frequência de um gene A devida à mutação de A para a em uma taxa constante (μ) de 10^{-5} .

A maioria das taxas de mutação que foram determinadas são a soma de todas as mutações de A para *qualquer* forma mutante com um efeito detectável. O processo de mutação é ainda mais lento se considerarmos o aumento de *determinado* tipo alélico novo. Qualquer substituição *específica* de base é de, pelo menos, duas ordens de magnitude em frequência do que a soma de todas as mudanças.

Variação por recombinação

Quando uma nova mutação de um gene surge em uma população, ela ocorre como um evento simples em uma cópia particular de um cromossomo portado por algum indivíduo. Mas essa cópia cromossômica tem uma composição alélica particular para todos os outros genes polimórficos no cromossomo. Assim, se algum alelo mutante a surge no locus A de um cromossomo que já tem o alelo b no locus B , então, sem recombinação, todos os gametas levando o alelo a levariam também o alelo b em gerações futuras. A população então conteria apenas o haplótipo original AB e o novo haplótipo ab que surgiu da mutação. A recombinação entre o gene A e o gene B no duplo heterozigoto $AB/a b$, entretanto, produziria dois novos haplótipos $A b$ e $a B$.

A consequência da recombinação repetida entre genes é fazer combinações aleatórias de alelos de genes diferentes. Se a

frequência alélica de a no locus A é, digamos, 0,2 e a frequência do alelo b no locus B é, digamos, 0,4, então a frequência de ab seria $(0,2)(0,4) = 0,08$ se as combinações fossem aleatórias. Essa condição aleatória é o **equilíbrio de ligação**.

A recombinação entre genes no mesmo cromossomo não produz equilíbrio de ligação em uma única geração se os alelos em genes diferentes começam em associação não-aleatória uma com outra. Essa associação original, **desequilíbrio de ligação**, decai apenas lentamente de geração a geração a uma taxa que é proporcional à quantidade de recombinação entre os genes. Esse fato pode ser usado para encontrar a localização de genes desconhecidos em cromossomos, e para dar evidência de que alguma variante fenotípica é, de fato, influenciada por um gene desconhecido. Suponha que as pessoas que sofrem de alguma doença, digamos, diabetes, também levam um alelo de algum gene marcador que não tem nada a ver com a formação de insulina mais frequentemente do que seria esperado se a associação entre diabetes e o alelo marcador fosse aleatória. Esse achado evidenciaria que o diabetes é influenciado por um gene no mesmo cromossomo que aquele contendo o gene marcador e, se o desequilíbrio de ligação fosse muito forte, que o gene relacionado ao diabetes estivesse razoavelmente perto do marcador. A existência de tal desequilíbrio de ligação seria supostamente o resultado

acidental da origem mutacional do alelo marcador na mesma cópia cromossômica que ter o alelo associado ao diabetes.

A criação de variação genética por recombinação pode ser um processo muito mais rápido do que a criação por mutação. Essa alta taxa de reprodução de variação é simplesmente uma consequência de um número muito grande de diferentes recombinações cromossômicas que podem ser produzidas se levarmos em conta apenas crossings únicos. Se um par de cromossomos homólogos é heterozigoto em n loci, então um crossing pode ocorrer em qualquer um dos $n - 1$ intervalos entre eles, e, como cada recombinação produz dois produtos recombinantes, existem $2(n - 1)$ novos tipos gaméticos únicos de uma única geração de crossing over, mesmo considerando apenas crossings únicos. Se os loci heterozigotos estiverem bem espalhados ao longo do cromossomo, esses novos tipos gaméticos serão frequentes e uma considerável variação será gerada. Organismos assexuais ou organismos tais como bactérias que muito raramente sofrem recombinação sexual não têm essa fonte de variação; logo, novas mutações são o único meio pelo qual podem ser obtidas mudanças nas combinações dos genes. Como resultado, as populações de organismos assexuais podem mudar mais lentamente que as de organismos sexuais.

Variação por migração

Uma outra fonte de variação é a migração para uma população vinda de outra população com frequências gênicas diferentes. A população mista resultante terá uma frequência que é intermediária ao valor original e às frequências da população doadora.

Suponha que uma população recebe um grupo de migrantes e o número de imigrantes é igual a, digamos, 10% do tamanho da população nativa. Então, as populações mistas recém-formadas terão uma frequência alélica que é uma mistura 0,90:0,10 entre a frequência alélica original e a frequência alélica da população doadora. Se sua frequência alélica original de A fosse, digamos, 0,70, enquanto a população original tivesse uma frequência alélica A apenas de, digamos, 0,40, a nova população mista teria uma frequência de A que era $(0,70 \times 0,90) + (0,40 \times 0,10) = 0,67$. O Boxe 17.4 deriva o resultado geral. Como mostrado no Boxe 17.4, a mudança na frequência gênica é proporcional à diferença na frequência entre a população receptora e a média das populações doadoras. Ao contrário da taxa de mutação, a taxa de migração (m) pode ser grande; assim, se a diferença na frequência alélica

entre a população doadora e a população receptora for grande, a mudança em frequência pode ser substancial.

Devemos compreender *migração* como significando qualquer forma de introdução de genes de uma população na outra. Assim, por exemplo, os genes europeus “migraram” para uma população de origem africana na América do Norte constantemente desde que os africanos foram introduzidos como escravos. Podemos determinar a quantidade dessa migração observando a frequência de um alelo que é encontrado apenas em europeus e não em africanos, e comparando sua frequência entre negros na América do Norte. Podemos usar a fórmula da mudança de frequência gênica da migração se a modificarmos ligeiramente para levar em conta o fato de que ocorreram várias gerações de mistura. Se a taxa de mistura não for muito grande, então (em uma ordem menor de aproximação) a soma das taxas de migração da única geração por várias gerações (chamemos isso de M) estará relacionada à mudança total na população receptora após essas várias gerações pela mesma expressão que a usada no Boxe 17.4 para mudanças devidas a uma única geração de migração. Se, como antes, P é a frequência alélica da população doadora e p_0 é a frequência original entre os receptores, então, na população receptora,

$$\Delta p_{\text{total}} = M(P - p_0)$$

Logo,

$$M = \frac{\Delta p_{\text{total}}}{P - p_0}$$

Por exemplo, o alelo Fy^a do grupo sanguíneo Duffy está ausente na África, mas tem uma frequência de 0,42 nos brancos do estado da Geórgia. Entre os negros da Geórgia, a frequência de Fy^a é de 0,046. Portanto, a migração total de genes da população de brancos para os negros desde a introdução dos escravos no século XVIII é

$$M = \frac{\Delta p_{\text{total}}}{P - p_0} = \frac{(0,046 - 0,0)}{(0,42 - 0,0)} = 0,1095$$

Isto é, na média, em todos os americanos com ancestrais africanos na Geórgia, cerca de 11% de seus alelos gênicos foram derivados de um ancestral europeu. Essa percentagem, entretanto, é apenas uma média, e diferentes americanos identificados como “negros” têm proporções diferentes de ancestrais europeus e africanos. Quando a mesma análise é feita em negros americanos de Oakland (Califórnia) e Detroit, M é 0,22 e 0,26,

Boxe 17.4

O Efeito da Migração na Frequência Alélica

Se p_t é a frequência de um alelo em uma população receptora na geração t , P é a frequência desse alelo em uma população doadora (ou a média de várias populações doadoras) e m é a proporção da população receptora que é feita de novos migrantes da população doadora, então a frequência alélica na população receptora na geração seguinte, p_{t+1} , é o resultado da mistura de genes $1 - m$ da

população receptora com m genes da população doadora. Assim,

$$p_{t+1} = (1 - m)p + mP = p_t + m(P - p_t)$$

e

$$\Delta p = p_{t+1} - p_t = m(P - p_t)$$

respectivamente, mostrando taxas maiores de mistura nessas cidades do que na Geórgia, ou um movimento diferencial nessas cidades por negros americanos que têm mais ancestrais europeus. Em qualquer caso, a variação genética do locus *Fy* foi aumentada por essa mistura. Ao mesmo tempo, a frequência da mutação *Hb^S* da anemia falciforme foi diminuída em afro-americanos entre 10 e 20% desse valor em suas populações africanas ancestrais como resultado da mistura.

► 17.4 Seleção

Até agora neste capítulo, consideramos as mudanças em uma população que surgem das forças da mutação, migração, recombinação e estrutura reprodutiva. Mas essas mudanças não podem explicar por que os organismos parecem tão bem adaptados a seus ambientes, já que eles são aleatórios com relação ao modo pelo qual cada organismo vive em seu ambiente. As mudanças em uma espécie em resposta a mudanças ambientais ocorrem devido a genótipos diferentes, produzidos por mutação e recombinação, terem habilidades diferentes em sobreviver e se reproduzir. As taxas diferenciais e sobrevivência e reprodução são o significado da **seleção**, e o processo de seleção altera as frequências de vários genótipos na população. Darwin chamou esse processo de sobrevivência diferencial e reprodução de tipos diferentes de **seleção natural** por analogia com a **seleção artificial** feita pelos criadores de animais e plantas, quando deliberadamente selecionam alguns indivíduos de um tipo preferido.

A probabilidade relativa de sobrevivência e taxa de reprodução de um fenótipo ou genótipo hoje é chamada de **adaptabilidade darwiniana**. Embora os geneticistas às vezes falem genericamente da adaptabilidade de um indivíduo, o conceito de adaptabilidade de fato aplica-se à probabilidade média de sobrevivência de membros individuais de uma classe fenotípica ou genotípica. Devido a eventos casuais nas histórias de vida dos indivíduos, mesmo dois organismos com genótipos idênticos, vivendo em ambientes idênticos, não vivem até a mesma idade ou deixam o mesmo número de prole. É a adaptabilidade de um genótipo em média de todos os que o possuem que importa.

Adaptabilidade é uma consequência da relação entre o fenótipo de um organismo e o ambiente no qual o organismo vive. O mesmo genótipo terá adaptabilidades diferentes em ambientes diferentes. Um motivo é que mesmo organismos geneticamente idênticos podem desenvolver fenótipos diferentes se expostos a ambientes diferentes durante o desenvolvimento. Mas, ainda que o fenótipo seja o mesmo, o sucesso do organismo depende do ambiente. Ter pés alados é muito bom para nadar, mas é uma desvantagem para andar no chão, como revelará a observação por alguns minutos de como os patos andam. Nenhum genótipo é incondicionalmente superior em adaptabilidade a todos os outros em todos os ambientes.

A adaptabilidade reprodutiva não deve ser confundida com “adaptabilidade física” no sentido usual do termo, embora ambos os termos possam estar relacionados. A despeito de quão saudável, forte e mentalmente ativo possa ser o possuidor de um genótipo, esse genótipo tem uma adaptabilidade

zero se, por algum motivo, seus possuidores não deixarem prole. A adaptabilidade de um genótipo é uma consequência de todos os efeitos fenotípicos dos genes envolvidos. Assim, um alelo que dobra a fecundidade de seus portadores mas, ao mesmo tempo, reduz o tempo médio de vida de seus possuidores em 10% será mais adaptado que seus alternativos, a despeito de sua propriedade de vida curta. O exemplo mais comum é o cuidado parental. Uma ave adulta que gasta muita energia à procura de alimentos para seus filhotes terá uma probabilidade menor de sobrevivência do que a que fica com todo o alimento para si. Mas uma ave totalmente egoísta não deixará prole, pois seus filhotes não podem se alimentar por si mesmos. Como consequência, os cuidados parentais são favorecidos pela seleção natural.

Duas formas de seleção

Como as diferenças na reprodução e sobrevivência entre genótipos dependem do ambiente no qual os genótipos vivem e se desenvolvem, e porque os organismos podem alterar seus próprios ambientes, existem duas formas fundamentalmente diferentes de seleção. No caso simples, a adaptabilidade de um indivíduo não depende da composição da população à qual ele pertence; ela é uma propriedade fixa do fenótipo do indivíduo e do ambiente físico externo. Por exemplo, a habilidade relativa de duas plantas que vivem nas margens de um deserto para conseguir água suficiente irá depender da profundidade de suas raízes e de quanta água elas perdem pela superfície de suas folhas. Essas características são uma consequência de seus padrões desenvolvimentais e não são sensíveis à composição da população na qual vivem. A adaptabilidade de um genótipo em tal caso não depende do quanto é raro nem do quanto ele é frequente na população. A adaptabilidade é então **independente da frequência**.

Em contraste, considere organismos que estão competindo para pegar uma presa ou para evitar ser capturados por um predador. Então as abundâncias relativas de dois genótipos diferentes afetarão suas relativas adaptabilidades. Um exemplo é o mimetismo mülleriano em borboletas. Algumas espécies de borboletas com cores brilhantes (tais como as monarcas e vice-reis) são desagradáveis aos pássaros, que aprendem, após algumas tentativas, a evitar atacar borboletas com um padrão que se associe ao gosto ruim. Dentro de uma espécie que tem mais de um padrão, os padrões mais raros serão selecionados contra. Quanto mais raro o padrão, maior a desvantagem seletiva, porque os pássaros, muito provavelmente, terão tido uma experiência anterior de um padrão de baixa frequência e, portanto, não o evitarão. Essa seleção de se misturar na multidão é um exemplo de **adaptabilidade dependente de frequência**, pois a adaptabilidade de um tipo muda à medida que ele se torna mais ou menos frequente em uma população.

Por razões de conveniência matemática, a maioria dos modelos de seleção natural são baseados em adaptabilidade independente de frequência. Na verdade, porém, um número muito grande de processos seletivos (talvez a maioria) são dependentes de frequência. A cinética de mudança de frequência alélica depende da forma exata da dependência de

freqüência, e, apenas por esse motivo, torna-se difícil fazer qualquer generalização. Por motivo de simplicidade e como ilustração das principais características qualitativas da seleção, lidamos só com modelos de seleção independentes de freqüência neste capítulo, mas conveniência não deve ser confundida com realidade.

Medindo as diferenças de adaptabilidade

Na maior parte, podemos medir a adaptabilidade diferencial de genótipos diferentes mais facilmente quando os genótipos diferem em muitos loci. Em poucos casos, como os mutantes de laboratório, variedades hortícolas e grandes distúrbios metabólicos, uma substituição alélica em um único locus faz uma grande diferença no fenótipo para alterar mensuravelmente a adaptabilidade. A Figura 17.8 mostra a probabilidade de sobrevivência do zigoto ao adulto, isto é, a **viabilidade**, em três temperaturas diferentes de várias linhagens diferentes tornadas homozigotas para o segundo cromossomo de *D. pseudoobscura*. Esses cromossomos foram amostrados de uma população natural e levaram uma variedade de alelos diferentes em loci diferentes, como esperávamos da grande quantidade de variação de nucleotídeos presentes na natureza (veja anteriormente). Como geralmente é o caso, a adaptabilidade (nesse caso, um componente da adaptabilidade total, viabilidade) é diferente em ambientes diversos.

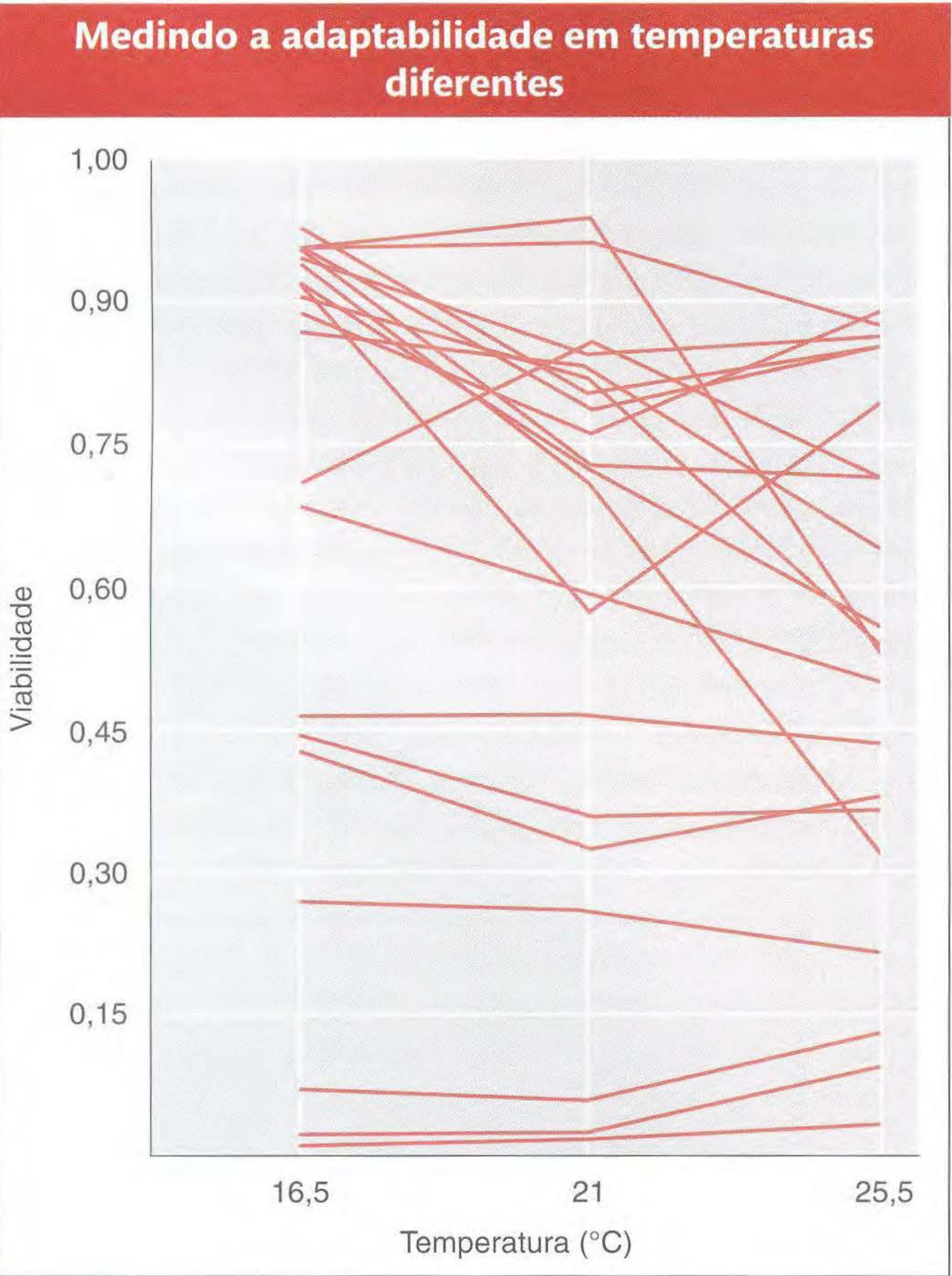


FIG. 17.8 Viabilidades de vários cromossomos homozigotos de *Drosophila pseudoobscura* em três temperaturas diferentes.

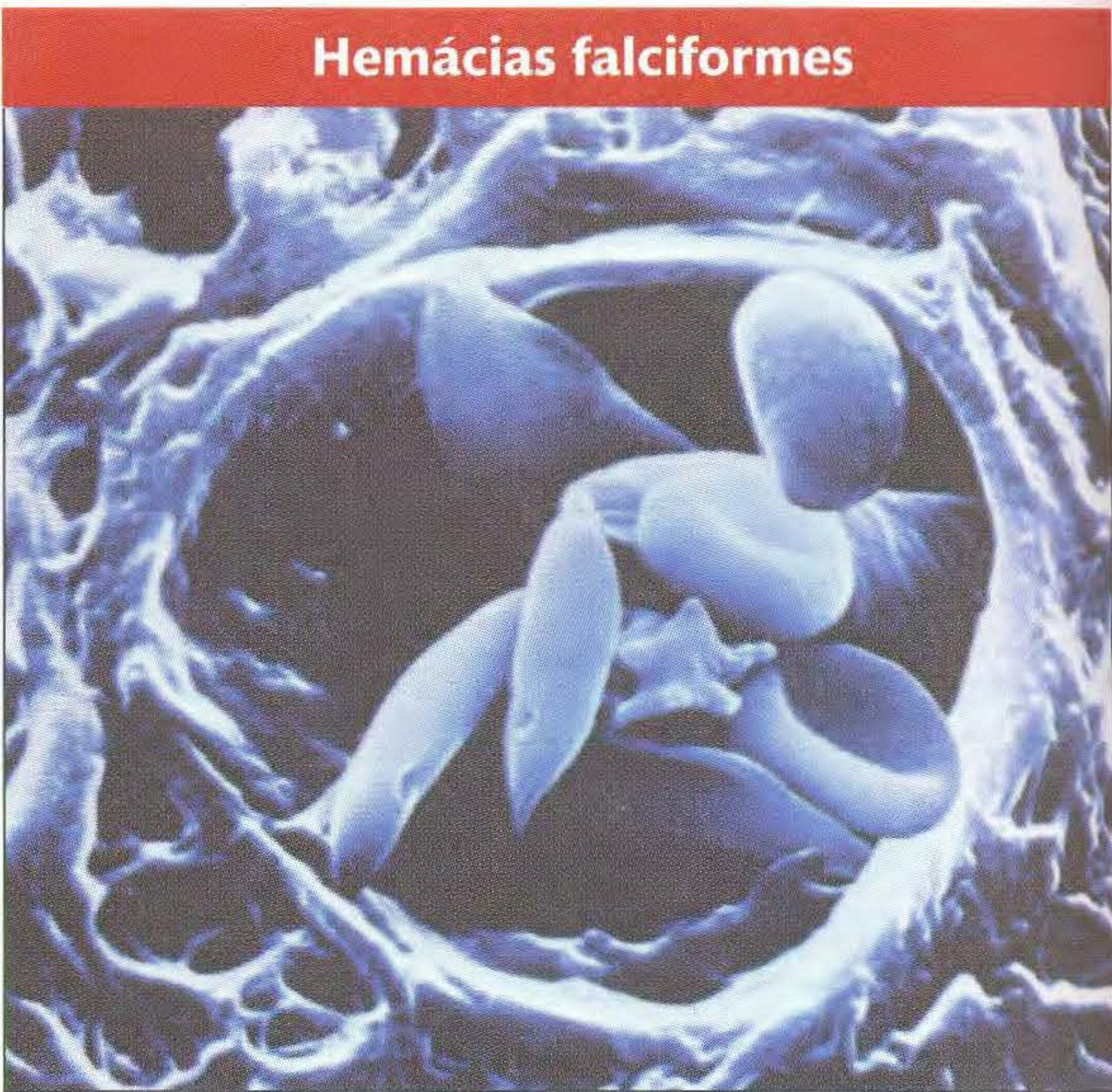


FIG. 17.9 Hemácias de uma pessoa com anemia falciforme. Algumas hemácias normais discóides são circundadas por células distorcidas falciformes.

O estado homozigoto é letal ou quase em alguns casos em todas as três temperaturas, enquanto alguns casos têm consistentemente alta variabilidade. A maioria dos genótipos, entretanto, não é consistente em viabilidade entre as temperaturas, e nenhum genótipo é incondicionalmente o mais adaptado a todas as temperaturas.

Existem casos nos quais as substituições de um só gene levam a nítidas diferenças de adaptabilidade. São exemplos os muitos “erros inatos do metabolismo”, nos quais um alelo recessivo interfere na via metabólica e é letal em homozigotos. Um exemplo é a anemia falciforme. As pessoas com esse distúrbio são homozigotas para o alelo que codifica a hemoglobina S em vez da hemoglobina normal. Eles morrem de uma grave anemia porque a hemoglobina S cristaliza em baixas concentrações de oxigênio, fazendo com que as hemácias se tornem falciformes e se rompam (Figura 17.9).

Como descrito no Capítulo 6, um outro exemplo em humanos é a fenilcetonúria, na qual o tecido degenera como resultado do acúmulo de um intermediário tóxico na via de metabolismo da tirosina. Esse caso também ilustra como a adaptabilidade é alterada por mudanças no ambiente. As pessoas que nascem com PKU sobreviverão se observarem uma dieta restrita isenta de tirosina.

Como a seleção funciona

A seleção atua alterando as freqüências alélicas em uma população. O modo mais simples de ver o efeito da seleção é considerar um alelo *a* que é completamente letal antes da idade reprodutiva na condição homozigota, tal como o alelo que leva à doença de Tay-Sachs. Suponha que, em alguma geração, a freqüência desse gene é 0,10. Então, em uma população de reprodução aleatória, as proporções dos três genótipos após a fertilização são

A/A	A/a	a/a
0,81	0,18	0,01

Na idade reprodutiva, entretanto, os homozigotos a/a já terão morrido, deixando os genótipos nesse estágio como

A/A	A/a	a/a
0,81	0,18	0,00

Mas essas proporções somam apenas 0,99 porque apenas 99% da população ainda sobrevive. Entre a população reprodutiva que atualmente sobrevive, a proporção deve ser recalculada dividindo-se por 0,99, de modo que as proporções totais somam 1,00. Após esse reajuste, temos

A/A	A/a	a/a
0,818	0,182	0,00

A frequência do alelo letal a entre os gametas produzidos por esses sobreviventes é então

$$0,00 + 0,182/2 = 0,091$$

e a mudança na frequência do alelo letal a em uma geração, expressa como o novo valor menos o antigo, foi $0,091 - 0,100 = -0,009$. Contrariamente, a mudança na frequência do alelo normal foi $+0,009$. Podemos repetir esse cálculo em cada geração sucessiva para obter as frequências previstas de alelos letais e normais em uma sucessão de gerações futuras.

O mesmo tipo de cálculo pode ser feito se os genótipos não forem simplesmente letais ou normais, mas cada genótipo tiver alguma probabilidade de sobrevivência. Esse cálculo geral é mostrado no Boxe 17.5. Após uma geração de seleção, o novo valor da frequência de A é igual ao antigo valor, p , multiplicado pela proporção da adaptabilidade média dos alelos A , \bar{W}_A , em relação à adaptabilidade de toda a população, \bar{W} . Se a adaptabilidade dos alelos A for maior que a adaptabilidade média de todos os alelos, então \bar{W}_A/\bar{W} é maior do que a unidade, e p' , o novo valor de frequência gênica após uma geração, é maior do que o antigo valor p . Assim, o alelo A aumenta na população. Contrariamente, se \bar{W}_A/\bar{W} é menor que a unidade, A diminui. Mas a adaptabilidade média da população, \bar{W} , é a média da adaptabilidade dos alelos A e os alelos a . Assim, se \bar{W}_A , a adaptabilidade média dos alelos A , é maior que a adaptabilidade média da população, ela também deve ser maior que \bar{W}_a , a adaptabilidade média dos alelos a . Assim, o alelo com a mais alta adaptabilidade média aumenta de frequência.

Note que a adaptabilidade $W_{A/A}$, $W_{A/a}$ e $W_{a/a}$ pode ser expressa como probabilidades absolutas de sobrevivência e taxas absolutas de reprodução, ou podem ser todas refeitas com relação a uma das adaptabilidades, que é dada como o valor padrão de 1,0. Essa reestruturação não tem nenhum efeito absoluto na fórmula de p' , pois cancela o numerador e o denominador.

Como resultado da seleção, o alelo com a mais alta adaptabilidade média relativa à adaptabilidade média de outros alelos aumenta de frequência na população.

Um aumento do alelo com a adaptabilidade mais alta significa que a adaptabilidade média da população como um todo aumenta, e, assim, a seleção também pode ser descrita

como um processo de *aumento da adaptabilidade média*. Essa regra é estritamente verdadeira apenas para adaptabilidade genotípica independente de frequência, mas é bem próxima da regra geral para ser usada como uma generalização proveitosa. Essa maximização da adaptabilidade não leva necessariamente a nenhuma propriedade ótima para a espécie como um todo, pois a adaptabilidade é definida apenas em relação a uma outra dentro de uma população. É a adaptabilidade relativa (não a absoluta) que está aumentada pela seleção. A população nem se torna maior ou cresce mais rápido nem tem menos probabilidade de se extinguir. Por exemplo, suponha que um alelo faz com que seus portadores ponham mais ovos do que outros genótipos na população. Esse alelo de maior fecundidade aumentará na população. Mas o tamanho da população no estágio adulto depende do total de alimentos disponíveis nos estágios imaturos; assim, não haverá um aumento no tamanho total da população, mas apenas um aumento no número de indivíduos imaturos que ficam em inanição até morrer antes da vida adulta.

Taxa de mudança em frequência gênica

A expressão geral de mudança na frequência alélica derivada no Boxe 17.5 é particularmente esclarecedora. Ela diz que Δp será positivo (A não aumentará) se a adaptabilidade média dos alelos A for maior que a adaptabilidade média dos alelos a , como já vimos. Mas também mostra que a velocidade depende não só da diferença em adaptabilidade entre os alelos, mas também do fator pq , que é proporcional à frequência de heterozigotos ($2pq$). Para determinada diferença na adaptabilidade dos alelos, sua frequência mudará mais rapidamente quando os alelos A e a estão em frequência intermediária, e, assim, pq é grande. Se p estiver perto de 0 ou 1 (isto é, se A ou a estiver quase fixado na frequência 0 ou 1), então pq é quase 0 e a seleção continuará bem lentamente.

A curva em forma de S na Figura 17.10 representa o curso da seleção de um novo alelo favorável A que, recentemente, entrou em uma população de homozigotos a/a . No início, a mudança na frequência é muito pequena porque p ainda está perto de 0. Então ela se acelera à medida que A se torna mais frequente, mas diminui novamente à medida que A sobe e a

Um novo alelo favorável A aumenta de frequência rapidamente à medida que se torna mais comum



FIG. 17.10 O padrão de tempo do aumento de frequência de um novo alelo favorável A que entrou em uma população de homozigotos a/a .

Boxe 17.5 O Efeito da Seleção nas Frequências Alélicas

Suponha que uma população está se reproduzindo aleatoriamente com relação a determinado locus com dois alelos, e que a população é tão grande que (por enquanto) podemos ignorar a endogamia. Logo após os ovócitos terem sido fertilizados, os genótipos dos zigotos estarão em equilíbrio de Hardy-Weinberg:

Genótipo	A/A	A/a	a/a
Frequência	p^2	$2pq$	q^2

e

$$p^2 + 2pq + q^2 = (p + q)^2 = 1,0$$

onde p é a frequência de A .

Suponha ainda que os três genótipos têm as probabilidades relativas de sobrevivência até a vida adulta (viabilidades) de $W_{A/A}$, $W_{A/a}$ e $W_{a/a}$. Por simplificação, vamos supor que todas as diferenças de adaptabilidade são diferenças de sobrevivência entre os zigotos e o estágio adulto. (Diferenças em fertilidade originam formulações matemáticas muito mais complexas.) Entre a prole, após terem atingido a vida adulta, as frequências serão

Genótipo	A/A	A/a	a/a
Frequência	$p^2 W_{A/A}$	$2pq W_{A/a}$	$q^2 W_{a/a}$

Essas frequências ajustadas não se somam até a unidade, porque os W são todos frações menores que 1. Entretanto, podemos reajustá-las de modo que o façam, sem mudar sua relação uma para outra, dividindo cada frequência pela soma das frequências após a seleção (\bar{W}):

$$\bar{W} = p^2 W_{A/A} + 2pq W_{A/a} + q^2 W_{a/a}$$

Assim definido, \bar{W} é chamado de **adaptabilidade média** da população porque é, de fato, a média da adaptabilidade de todos os indivíduos na população. Após esse ajuste, temos

Genótipo	A/A	A/a	a/a
Frequência	$p^2 \frac{W_{A/A}}{\bar{W}}$	$2pq \frac{W_{A/a}}{\bar{W}}$	$q^2 \frac{W_{a/a}}{\bar{W}}$

Podemos agora determinar a frequência p' do alelo A na geração seguinte somando os genes:

$$p' = A/A + \frac{1}{2} A/a = p^2 \frac{W_{A/A}}{\bar{W}} + pq \frac{W_{A/a}}{\bar{W}} = p^2 \frac{W_{A/A}}{\bar{W}} + q \frac{W_{A/a}}{\bar{W}}$$

Finalmente, notamos que a expressão $pW_{A/A} + qW_{A/a}$ é a adaptabilidade média dos alelos A porque, pelas frequências de Hardy-Weinberg, uma proporção p de todos os alelos A estão presentes nos homozigotos com outro A , em cujo caso eles têm uma adaptabilidade de $W_{A/A}$, enquanto uma proporção q de todos os alelos A estão presentes em heterozigotos com a e têm uma adaptabilidade $W_{A/a}$. Usando \bar{W}_A para indicar $pW_{A/A} + qW_{A/a}$, notamos que a adaptabilidade média do alelo A produz a nova frequência alélica final

$$p' = p \frac{\bar{W}_A}{\bar{W}}$$

Um modo alternativo de ver o processo de seleção é resolver a *mudança* em frequências alélicas em uma geração:

$$\Delta p = p' - p = p \frac{\bar{W}_A}{\bar{W}} - p = \frac{p(\bar{W}_A - \bar{W})}{\bar{W}}$$

Mas \bar{W} , a adaptabilidade média da população, é a média das adaptabilidades alélicas \bar{W}_A e \bar{W}_a ; logo,

$$\bar{W} = p \bar{W}_A + q \bar{W}_a$$

onde \bar{W}_a é a adaptabilidade média dos alelos a . Substituindo essa expressão por \bar{W} na fórmula para Δp e lembrando que $q = 1 - p$, obtemos (após alguma manipulação algébrica)

$$\Delta p = \frac{pq(\bar{W}_A - \bar{W}_a)}{\bar{W}}$$

se torna muito raro (q fica perto de 0), o que é exatamente o esperado de um processo de seleção. Quando a maioria da população é de um tipo, não há nada para selecionar. Para que ocorra mudança por seleção natural, deve haver variação genética; quanto mais variação, mais rápido o processo.

Uma consequência da dinâmica mostrada na Figura 17.10 é que é extremamente difícil reduzir significativamente a frequência de um alelo que já é raro em uma população. Assim, os programas eugênicos destinados a eliminar alelos recessivos deletérios de populações humanas, evitando a reprodução de pessoas afetadas, não funciona. Logicamente, se todos os heterozigotos fossem impedidos de se

reproduzir, o alelo poderia ser eliminado (exceto por mutações novas) em uma única geração. Como cada ser humano é heterozigoto para vários alelos deletérios, entretanto, nenhum poderia se reproduzir.

Quando alelos alternativos não são raros, a seleção pode causar mudanças bem rápidas em frequência alélica. A Figura 17.11 mostra o curso da eliminação de um alelo de desidrogenase málica que tinha uma frequência inicial de 0,5 em uma população de laboratório de *D. melanogaster*. As adaptabilidades nesse caso são

$$W_{A/A} = 1,0 \quad W_{A/a} = 0,75 \quad W_{a/a} = 0,40$$

Um alelo menos favorável diminui de frequência

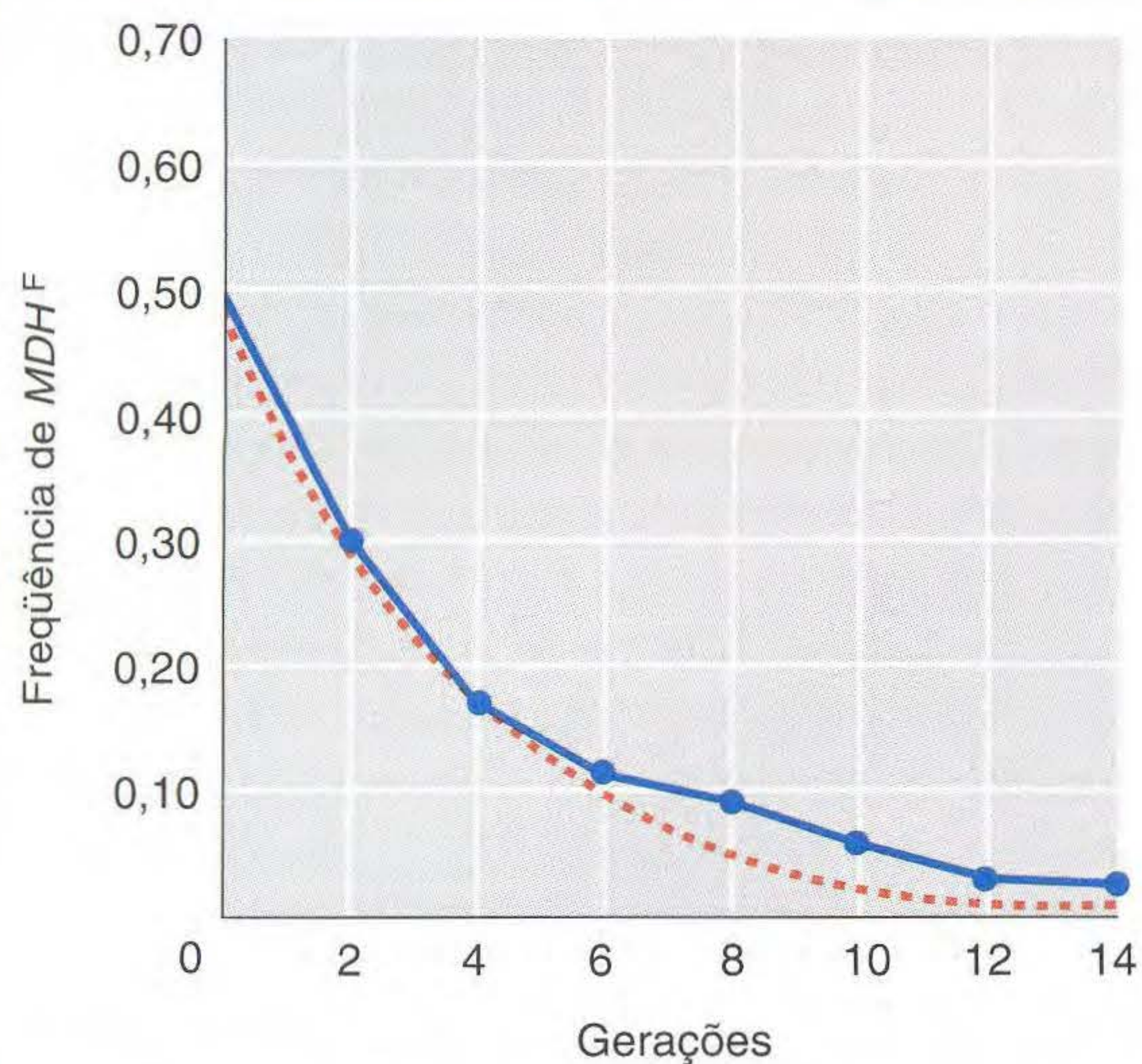


FIG. 17.11 A perda de um alelo MDH^F no locus da desidrogenase málica devido à seleção em uma população de laboratório de *Drosophila melanogaster*. A linha vermelha pontilhada mostra a curva teórica de mudança computada quanto à adaptabilidade de $W_{A/A} = 1,0$, $W_{A/a} = 0,75$ e $W_{a/a} = 0,4$. [De R.C. Lewontin, *The Genetic Basis of Evolutionary Change*. Copyright 1974 by Columbia University Press. Dados por cortesia de E. Berger.]

A frequência de a declina rapidamente, mas não é reduzida a 0, e, para continuar reduzindo sua frequência, seriam necessários tempos muito longos, como mostrado no caso da eugenia negativa.

A menos que os alelos alternativos estejam presentes em frequências intermediárias, a seleção (especialmente contra recessivos) é bem lenta. A seleção depende da variação genética.

► 17.5 Polimorfismo Balanceado

Até agora, consideramos as mudanças em frequências alélicas que ocorrem quando os portadores homozigotos de um alelo, digamos A/A , são mais adaptáveis que os portadores homozigotos de outro alelo, digamos a/a , enquanto os heterozigotos, A/a , estão entre A/A e a/a em suas adaptabilidades. Mas existem outras possibilidades.

Sobredominância e subdominância

Primeiro, o heterozigoto pode ser *mais* adaptado que ambos os homozigotos, uma condição chamada de **sobredominância** (*overdominance*) em adaptabilidade. Quando um dos alelos, digamos A , está em baixa frequência, praticamente não há homozigotos A/A , e o alelo está presente quase totalmente na condição heterozigota. Como os heterozigotos são mais adaptáveis que os homozigotos, quase todos os alelos A são levados pelo genótipo mais adaptável, e, assim, A aumentará em frequência enquanto a diminuirá em frequência. Por outro lado, quando a está em frequência muito baixa, ele está

presente quase que inteiramente na condição heterozigota. Nesse caso, quase todos os alelos a são levados pelo genótipo mais adaptável, e aumentarão de frequência à custa dos alelos A . O efeito final dessas duas pressões — uma aumentando a frequência de alelos A , quando são raros, e a outra aumentando os alelos a , quando *eles* são raros — é levar as frequências alélicas a um estado de equilíbrio estável que é intermediário às dos dois alelos. Qualquer desvio casual das frequências alélicas de um lado ou do outro do equilíbrio será contrabalançada pela força da seleção. Podemos simbolizar as adaptabilidades dos três genótipos por

$$\begin{array}{ccc} W_{A/A} & W_{A/a} & W_{a/a} \\ 1 - t & 1 & 1 - s \end{array}$$

onde t e s são desvantagens seletivas dos dois homozigotos. Então a frequência de equilíbrio do alelo A é simplesmente a proporção

$$p(A) = \frac{s}{s + t}$$

(Como um exercício avançado, você pode obter esse resultado notando que, no equilíbrio, a adaptabilidade média do alelo A , \bar{W}_A , é igual à adaptabilidade média do alelo a , \bar{W}_a . Ajustando essas adaptabilidades médias iguais uma à outra e resolvendo $p(A)$, temos o resultado.) Esse equilíbrio explica a alta frequência da condição anemia falciforme no Oeste da África. Os homozigotos para o alelo anormal Hb^S morrem prematuramente de anemia falciforme. Mas há uma alta mortalidade no Oeste da África pela malária falcípara, que mata muitos dos homozigotos para o alelo normal, Hb^A . Os heterozigotos, Hb^A/Hb^S , sofrem apenas de uma anemia branda, não-fatal, e eles estão protegidos contra a malária falcípara pela presença da hemoglobina anormal em suas hemácias. Essa sobredominância em adaptabilidade foi perdida, entretanto, quando foram trazidos escravos para o Novo Mundo porque a forma falcípara de malária não existia no hemisfério ocidental. Como consequência, entre os escravos e seus descendentes, a seleção foi apenas contra os homozigotos Hb^S/Hb^S , levando a uma redução na frequência desse alelo pela mortalidade. Recentemente, a anemia falciforme recebeu suficiente atenção médica para não ser mais uma fonte significativa de mortalidade, e, assim, a seleção contra o alelo Hb^S não é mais tão poderosa. Outras reduções na frequência do alelo serão então principalmente a consequência de mistura continuada com populações de ancestrais não-africanos.

Outra possível relação de adaptabilidade entre alelos é que o heterozigoto é menos adaptado que ambos os homozigotos (**subdominância** em adaptabilidade). Nesse caso, a seleção favorece um alelo que é comum, não quando ele é raro. Assim, uma frequência alélica intermediária é instável, e a população deve tornar-se fixada para o alelo A ou para o alelo a . O polimorfismo resultante de uma mistura de uma população A/A com uma população a/a deve ser rapidamente perdido. Um exemplo bem conhecido, mas misterioso, de subdominância em adaptabilidade é a incompatibilidade de Rh em humanos. As crianças Rh-positivas nascidas de mães Rh-negativas em geral sofrem de anemia hemolítica, quando

neonatos, porque suas mães produzem anticorpos contra as células sangüíneas do feto. As mães Rh-negativas são homozigotas, Rh^-/Rh^- , e, assim, sua prole Rh-positiva que morre de anemia deve ser heterozigota, Rh^-/Rh^+ . O mistério é que todas as populações humanas são polimórficas para os dois alelos Rh. Assim, esse polimorfismo humano deve ser muito antigo, antecedendo as raças geográficas modernas. A previsão teórica simples é que tal polimorfismo é instável e deve ter desaparecido.

Balanço entre mutação e seleção

O balanço de sobredominância de forças seletivas não é a única situação na qual um equilíbrio de frequências alélicas pode surgir. As frequências alélicas também podem alcançar o equilíbrio nas populações quando a introdução de novos alelos por mutação repetida é balanceado por sua remoção pela seleção natural. Esse balanço provavelmente explica a persistência de doenças genéticas como polimorfismos de baixo nível em populações humanas. Novas mutações deletérias estão constantemente surgindo espontaneamente ou como resultado da ação de mutágenos. Essas mutações podem ser completamente recessivas ou parcialmente dominantes. A seleção remove-as da população, mas há um equilíbrio entre o surgimento e a remoção.

A equação geral para esse equilíbrio é dada em detalhe no Boxe 17.6. Ela mostra que a frequência do alelo deletério em equilíbrio depende da proporção μ/s , onde μ é a possibilidade de uma mutação ocorrer em um gameta recém-formado (taxa de mutação), e s é a intensidade de seleção contra o genótipo deletério. Para um alelo deletério totalmente recessivo cuja adaptabilidade no estado homozigoto é $1 - s$, a frequência de equilíbrio é

$$q = \sqrt{\frac{\mu}{s}}$$

Assim, por exemplo, um letal recessivo ($s = 1$) mutando a uma taxa de $\mu = 10^{-6}$ terá uma frequência de equilíbrio de 10^{-3} . De fato, se soubermos que um alelo era um letal recessivo e não tinha efeitos quando presente em uma única cópia em

heterozigotos, podemos estimar sua taxa de mutação como o quadrado de sua frequência. Mas a base biológica das suposições antecedentes a tais cálculos devem ser firmes. A anemia falciforme já foi considerada letal recessiva sem efeitos nos heterozigotos. Essa suposição levou a uma taxa de mutação estimada na África de 0,1 para esse locus como uma explicação da alta frequência de equilíbrio de tal letal. Entretanto, hoje sabemos que seu equilíbrio resulta do fato de os heterozigotos terem uma adaptabilidade maior que a dos indivíduos normais ou dos homozigotos mutantes, resultando em um equilíbrio estável das duas formas alélicas. Também podemos obter o equilíbrio entre a seleção e a mutação para um alelo deletério que tem algum efeito deletério nos heterozigotos, bem como seu efeito nos homozigotos. Se fizermos a adaptabilidade ser $W_{A/A} = 1,0$, $W_{A/a} = 1 - hs$ e $W_{a/a} = 1 - s$ para um alelo parcialmente dominante a , onde h é o grau de dominância do alelo deletério, então um cálculo similar ao do Boxe 17.6 nos dará

$$\hat{q} = \frac{\mu}{hs}$$

Assim, se $\mu = 10^{-6}$ e o letal não for totalmente recessivo mas tiver 5% de efeito deletério nos heterozigotos ($s = 1,0$, $h = 0,05$), então

$$\hat{q} = \frac{10^{-6}}{0,05} = 2 \times 10^{-5}$$

que é menor em duas ordens de grandeza que a frequência de equilíbrio para o caso puramente recessivo. Em geral, podemos esperar alelos deletérios, completamente recessivos, tendo frequências muito mais altas que os parcialmente dominantes, pois os alelos recessivos estão protegidos nos heterozigotos.

► 17.6 Eventos Aleatórios

Se uma população consiste em um número finito de indivíduos (como são todas as populações reais) e determinado par de genitores tiver apenas uma prole pequena, então, mesmo na ausência de todas as forças seletivas, a frequência de um

Boxe 17.6 O Balanço Entre Seleção e Mutação

Se q for a frequência de um alelo deletério a e $p = 1 - q$ for a frequência do alelo normal A , então a mudança de frequências alélicas devida à taxa de mutação μ é

$$\Delta q_{\text{mut}} = \mu p$$

Um modo simples de expressar a adaptabilidade dos genótipos no caso de um alelo deletério recessivo a é $W_{A/A} = W_{Aa} = 1,0$ e $W_{a/a} = 1 - s$, onde s é a perda de adaptabilidade nos homozigotos recessivos. Agora podemos substituir essas adaptabilidades em nossa expressão geral para a mudança de frequência alélica (veja Boxe 17.5) e obter

$$\Delta q_{\text{sel}} = \frac{-pq(sq)}{1 - sq^2} = \frac{-spq^2}{1 - sq^2}$$

Equilíbrio significa que o aumento na frequência alélica devido à mutação deve balancear exatamente a diminuição da frequência alélica devida à seleção; logo,

$$\Delta \hat{q}_{\text{mut}} + \Delta \hat{q}_{\text{sel}} = 0$$

Lembrando que \hat{q} em equilíbrio será bem pequeno, $1 - s\hat{q}^2 \approx 1$, e substituindo os termos por $\Delta \hat{q}_{\text{mut}}$ e $\Delta \hat{q}_{\text{sel}}$ na fórmula anterior, teremos

$$\frac{\mu \hat{p} - s \hat{p} \hat{q}^2}{(1 - s \hat{q}^2)} \approx \mu \hat{p} - s \hat{p} \hat{q}^2 = 0$$

ou, no equilíbrio

$$\hat{q}^2 = \frac{\mu}{s} \quad \text{e} \quad \hat{q} = \sqrt{\frac{\mu}{s}}$$

gene não será exatamente reproduzida na geração seguinte, devido a erros de amostragem. Se, em uma população de 1.000 indivíduos, a frequência de a é 0,5 em uma geração, ela pode por acaso ser 0,493 ou 0,505 na geração seguinte devido à produção aleatória de um pouco mais ou um pouco menos prole de cada genótipo. Na segunda geração, há um outro erro de amostragem baseado na nova frequência gênica, e, assim, a frequência de a pode ir de 0,505 a 0,511 ou de volta a 0,498. Esse processo de flutuação aleatória continua geração após geração, sem força para levar a frequência de volta a seu estado inicial, pois a população não tem “memória genética” de seu estado há muitas gerações. Cada geração é um evento independente. Essa mudança aleatória de frequências alélicas é conhecida como **deriva genética**.

O resultado final da deriva genética é que a população ocasionalmente deriva para $p = 1$ ou $p = 0$. Após esse ponto, não é mais possível nenhuma mudança; a população tornou-se homozigota. Uma população diferente, isolada da primeira, também sofre essa **deriva genética aleatória**, mas pode tornar-se homozigota para o alelo A , enquanto a primeira população tornou-se homozigota para o alelo a . Com o passar do tempo, as populações isoladas divergem uma da outra, cada uma perdendo heterozigose. A variação originalmente presente *dentro* das populações agora aparece como uma variação *entre* as populações.

Uma forma de deriva genética ocorre quando um pequeno grupo se separa de uma população maior para fundar uma nova colônia. Essa “deriva aguda”, chamada de **efeito do fundador**, resulta de uma única geração de amostragem de um pequeno número de colonizadores da grande população original, seguida de várias gerações durante as quais a nova colônia permanece em pequeno número. Mesmo que a população cresça após algum tempo, ela continuaria a derivar, mas a uma velocidade menor. O efeito do fundador é provavelmente responsável pela praticamente completa falta de grupo sanguíneo B em americanos nativos, cujos ancestrais chegaram em pequenos números através do Estreito de Bering no final da última era do gelo, há cerca de 20.000 anos, mas cuja população ancestral no Noroeste da Ásia teve uma frequência intermediária de grupo B.

O processo de deriva genética deve ser familiar. Ele é, de fato, um outro modo de observar o efeito da endogamia em populações pequenas já discutidas. As populações descendentes de um número muito pequeno de ancestrais têm uma alta probabilidade de que todas as cópias de um alelo particular sejam idênticas por descendência de um único ancestral comum (veja a Figura 17.6). Seja visto como endogamia ou como amostra aleatória de genes, o efeito é o mesmo. As populações não reproduzem exatamente suas constituições genéticas; há um componente aleatório de mudança de frequências gênicas.

Um resultado de amostra aleatória é que a maioria das novas mutações, mesmo que não sejam selecionadas contra, nunca são bem-sucedidas em tornar-se parte da composição genética em longo prazo da população. Suponha que um único indivíduo seja heterozigoto para uma nova mutação. Há alguma chance de que o indivíduo em questão não tenha

nenhuma prole. Mesmo que tenha um filho(a), há uma chance de $1/2$ de que a nova mutação não seja transmitida para essa prole. Se o indivíduo tiver dois filhos, a probabilidade de que nenhum da prole tenha a nova mutação é $1/4$, e assim por diante. Suponha que a nova mutação seja bem-sucedidamente transmitida para a prole. Então a loteria é repetida na geração seguinte, e, novamente, o alelo pode ser perdido. De fato, se uma população tem o tamanho N , a chance de que uma nova mutação seja eventualmente perdida por acaso é $(2N - 1)/2N$. (Para uma derivação desse resultado, que está além do escopo deste livro, veja os Capítulos 2 e 3 do *Principles of Population Genetics*, 3rd ed., by D.L. Hartl and A.G. Clark, Sinauer Associates, 1997.) Mas, se a nova mutação não for perdida, então a única coisa que pode acontecer a ela em uma população finita é que seja inserida na população e fique fixada. Esse evento tem a probabilidade de $1/(2N)$. Na ausência da seleção, então, a história de uma população assemelha-se ao que mostra a Figura 17.12. Por algum período de tempo, ela é homozigota; então, aparece uma nova mutação. Na maioria dos casos, o novo alelo mutante será perdido imediatamente ou pouco após aparecer. Ocasionalmente, entretanto, um novo alelo mutante deriva na população, e torna-se homozigoto para um novo alelo. O processo então começa novamente.

Um exemplo marcante do efeito da deriva genética nas populações humanas é o da variação nas frequências das variantes de tamanho das repetições VNTR entre populações de índios da América do Sul, como se observa no Quadro 17.5. Para um VNTR, D14S1, os suruí são muito variáveis, mas os karitianas, que vivem há muitas milhas de distância na floresta tropical, são quase homozigotos para uma variante,



FIG. 17.12 O surgimento, perda e eventual incorporação de novas mutações na vida de uma população. Se a deriva genética aleatória não causa a perda de uma nova mutação, então ela pode fazer com que toda a população torne-se homozigota para a mutação (na ausência de seleção). Nesta ilustração, surgiram 10 mutações, das quais 9 (vermelhas na parte inferior do gráfico) aumentam ligeiramente em frequência e então morrem. Apenas a quarta mutação a ocorrer (linha azul) se espalha na população [De J. Crow and M. Kimura, *An Introduction to the Population Genetics Theory*. Copyright 1970 by Harper & Row.]

supostamente devido à deriva genética nessas muito pequenas populações isoladas. Para outras VNTR, D14S13, nenhuma das populações tornou-se homozigota, mas o padrão de frequências dos alelos é muito diferente nas duas populações.

Mesmo uma nova mutação que é seletivamente um pouco favorável geralmente será perdida nas poucas primeiras gerações após aparecer na população, sendo uma vítima da deriva genética. Se uma nova mutação tem uma vantagem seletiva de s no heterozigoto no qual aparece, então a chance é apenas de $2s$ de que a mutação seja bem-sucedida em se instalar na população. Assim, uma mutação que é 1% melhor em adapta-

bilidade do que o alelo padrão na população será perdida em 98% das vezes por deriva genética. É até mesmo possível para uma mutação levemente deletéria aumentar em frequência e tornar-se fixada em uma população por deriva.

As novas mutações podem tornar-se estabelecidas em uma população muito embora não sejam favorecidas pela seleção natural simplesmente por um processo de deriva genética. Mesmo novas mutações favoráveis frequentemente são perdidas, e uma mutação ligeiramente deletéria pode se instalar em uma população por deriva.

► Resumo

O estudo das mudanças dentro de uma população, ou genética de populações, está relacionado a mudanças herdáveis nas populações ou organismos de processos individuais subjacentes de herança e desenvolvimento. A genética de populações é o estudo da variação herdada e sua modificação no tempo e no espaço.

A variação herdada identificável dentro das populações pode ser estudada examinando-se as diferenças em seqüências de aminoácidos específicos de proteínas, ou mesmo pelo exame, mais recente, de diferenças nas seqüências de nucleotídeos dentro do DNA. Esses tipos de observações revelaram que há um considerável polimorfismo em muitos loci dentro de uma população. Uma medida dessa variação é a quantidade de heterozigose em uma população. Tipicamente, uma população é polimórfica para 25 a 33% de seus genes codificantes de proteínas, e um indivíduo é heterozigoto em cerca de 10% de tais loci. Quaisquer dois humanos diferem em cerca de 3 milhões de nucleotídeos. Os estudos populacionais mostraram que, em geral, as diferenças genéticas entre indivíduos dentro das raças humanas são muito maiores que as diferenças médias entre as raças.

A fonte principal de toda a variação é a mutação. Entretanto, dentro de uma população, a frequência quantitativa de genótipos específicos pode ser mudada por recombinação, imigração de genes, eventos mutacionais continuados e pelo acaso.

Uma propriedade da segregação mendeliana é que as reproduções aleatórias resultem em uma distribuição de

equilíbrio de genótipos após uma geração. Entretanto, se houver endogamia, a variação genética dentro de uma população é convertida em diferenças entre as populações, tornando cada população separada homozigota para um alelo escolhido aleatoriamente. Por outro lado, para a maioria das populações, é atingido um balanço entre endogamia, mutação de um alelo para outro e imigração.

Um alelo pode aumentar ou diminuir de frequência dentro de uma população pela seleção natural de genótipos com probabilidades mais altas de sobrevivência e reprodução. Em muitos casos, tais mudanças levam à homozigose em determinado locus. Por outro lado, o heterozigoto pode ser mais adaptável que ambos os homozigotos, levando a um polimorfismo balanceado.

Em geral, a variação genética é o resultado da interação de forças. Por exemplo, um mutante deletério pode nunca ser totalmente eliminado de uma população, porque a mutação continuará a reintroduzi-lo na população. A migração também pode reintroduzir alelos que foram eliminados por seleção natural.

A menos que alelos alternativos sejam intermediários em frequência, a seleção (especialmente contra recessivos) é muito lenta, precisando de muitas gerações. Em muitas populações, especialmente as de pequeno tamanho, novas mutações podem tornar-se estabelecidas, muito embora não sejam favorecidas por seleção natural, ou podem ser eliminadas, muito embora sejam favorecidas, simplesmente por um processo de deriva genética aleatória.

► Termos-chave

adaptabilidade darwiniana
adaptabilidade dependente de frequência
adaptabilidade independente de frequência
adaptabilidade média
alelo fixado
coeficiente de endogamia
deriva genética
deriva genética aleatória

desequilíbrio de ligação
distribuição de equilíbrio
efeito do fundador
endogamia
endogamia negativa
equilíbrio de Hardy-Weinberg
equilíbrio de ligação
exogamia forçada
frequência alélica
frequência genotípica

genética de populações
haplótipo
heterozigose
homozigose por descendência
número variável de repetições em tandem (VNTR)
polimorfismo
polimorfismo de nucleotídeo único (SNP)
população

reprodução preferencial negativa
reprodução preferencial positiva
seleção
seleção artificial

seleção natural
sobredominância
subdominância
taxa de mutação

tipo selvagem
viabilidade

Problemas Resolvidos

Problema resolvido 1. Os polimorfismos para cor de conchas (amarelo ou rosa) e para a presença ou ausência de bandas nas conchas de *Cepaea nemoralis* são cada um o resultado de um par de alelos segregantes em um locus separado. Crie um programa experimental que revele as forças que determinam a frequência e distribuição geográfica desses polimorfismos.

SOLUÇÃO

a. Descreva as frequências de morfos diferentes para amostras de caramujos de um grande número de populações cobrindo uma faixa geográfica e ecológica de espécies. Cada caramujo deve ser anotado quanto a *ambos* os polimorfismos. Ao mesmo tempo, registre uma descrição do habitat de cada população. Avalie também o número de caramujos em cada população.

b. Faça a medida das distâncias de migração marcando uma amostra de caramujos com um ponto de tinta na concha, recolocando-o na população e, então, colhendo nova amostra posteriormente.

c. Crie uma porção de ovos postos por caramujos individuais de modo que os genótipos dos genitores masculinos possam ser deduzidos e os padrões reprodutivos não-aleatórios possam ser observados. As frequências de segregação *dentro* de cada família revelarão diferenças entre os genótipos na probabilidade de sobrevivência dos estágios iniciais de desenvolvimento.

d. Procure outras evidências de seleção de (1) padrões geográficos nas frequências dos alelos, (2) correlação entre frequências de alelos e variáveis ecológicas, incluindo densidade de população, (3) correlação entre as frequências de dois polimorfismos diferentes (populações com, digamos, altas frequências de conchas rosa também são caracterizadas por altas frequências de conchas com bandas) e (4) associações não-aleatórias *dentro* das populações de alelos em dois loci, indicando que algumas combinações podem ter uma adaptabilidade maior.

e. Procure evidências da importância da deriva genética aleatória comparando a variação nas frequências alélicas entre populações pequenas com a variação entre grandes populações. Se as pequenas populações podem variar mais de uma para outra do que as grandes, a deriva genética está implicada.

Problema resolvido 2. Cerca de 70% de todos os brancos norte-americanos sentem o gosto da feniltiocarbamida, e o restante não. A habilidade em sentir essa substância é determinada pelo alelo dominante *T*, e a incapacidade em sentir é determinada pelo alelo recessivo *t*. Se a popula-

ção supostamente está em equilíbrio de Hardy-Weinberg, quais são os genótipos e frequências alélicas nessa população?

SOLUÇÃO

Como 70% são sensíveis (*T/T* e *T/t*), 30% devem ser insensíveis (*t/t*). Essa frequência de homozigotos recessivos é igual a q^2 ; logo, para obtermos q , simplesmente tiramos a raiz quadrada de 0,30.

$$q = \sqrt{0,30} = 0,55$$

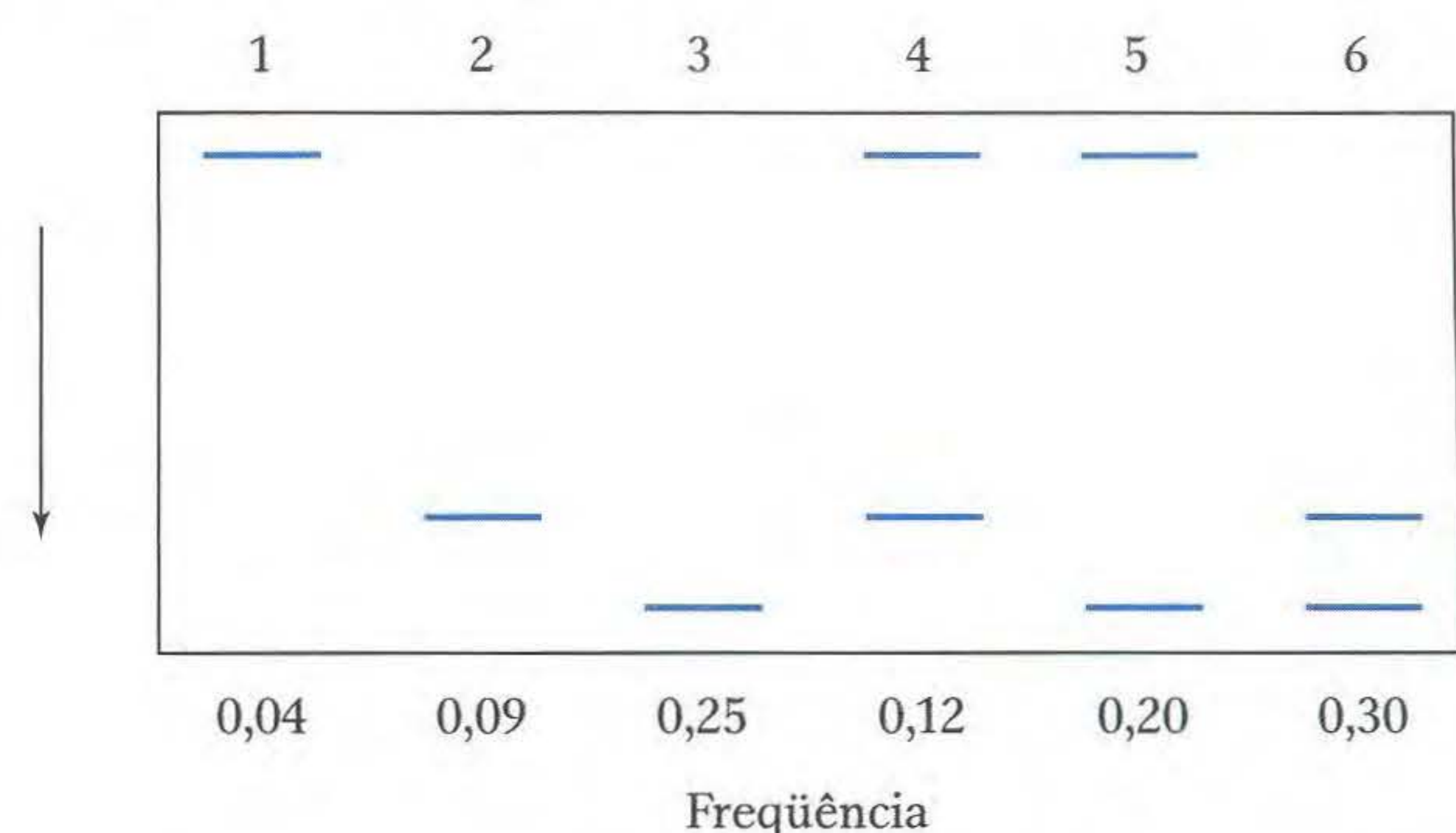
Como $p + q = 1$, podemos escrever

$$p = 1 - q = 1 - 0,55 = 0,45$$

Agora podemos calcular

$$\begin{aligned} p^2 &= (0,45)^2 = 0,20, \text{ a frequência de } T/T \\ 2pq &= 2 \times 0,45 \times 0,55 = 0,50, \text{ a frequência de } T/t \\ q^2 &= 0,3, \text{ a frequência de } t/t \end{aligned}$$

Problema resolvido 3. Em uma grande população natural de *Mimulus guttatus*, uma folha foi colhida de cada uma de um grande número de plantas. As folhas foram trituradas e submetidas a um gel de eletroforese. O gel foi então corado para uma enzima específica, X. Foram observados seis padrões de bandamento diferentes, como mostrado no diagrama seguinte.



- Supondo que esses padrões são produzidos por um único locus, proponha uma explicação genética para os seis tipos.
- Como você pode testar sua hipótese?
- Quais as frequências alélicas nessa população?
- A população está em equilíbrio de Hardy-Weinberg?

SOLUÇÃO

a. O exame do gel revela que existem apenas três posições de bandas: podemos chamá-las de lenta, intermediária e rápida, de acordo com a rapidez com que migraram no gel. Além disso, qualquer indivíduo pode apresentar uma banda ou

duas. A explicação mais simples é que existem três alelos de um locus (vamos chamá-los de S , I e F) e que os indivíduos com duas bandas são heterozigotos. Assim, a fileira 1 = S/S , 2 = I/I , 3 = F/F , 4 = S/I , 5 = S/F e 6 = I/F .

b. A hipótese pode ser testada fazendo cruzamentos controlados. Por exemplo, de uma autofecundação do tipo 5, podemos prever $\frac{1}{4} S/S$, $\frac{1}{2} S/F$ e $\frac{1}{4} F/F$.

c. As frequências podem ser calculadas por uma simples generalização de fórmulas dos dois alelos. Assim:

$$f_S = 0,04 + \frac{1}{2}(0,12) + \frac{1}{2}(0,20) = 0,20 = p$$

$$f_I = 0,09 + \frac{1}{2}(0,12) + \frac{1}{2}(0,30) = 0,30 = q$$

$$f_F = 0,25 + \frac{1}{2}(0,20) + \frac{1}{2}(0,30) = 0,50 = r$$

d. As frequências genotípicas de Hardy-Weinberg são

$$\begin{aligned}(p + q + r)^2 &= p^2 + q^2 + r^2 + 2pq + 2pr + 2qr \\ &= 0,04 + 0,09 + 0,25 + 0,12 + 0,20 + 0,30\end{aligned}$$

que são exatamente as frequências observadas. Assim, parece que a população está em equilíbrio.

Problema resolvido 4. Em uma grande população experimental de *Drosophila*, a adaptabilidade de um fenótipo recessivo é calculada como sendo 0,90, e a taxa de mutação do alelo recessivo é 5×10^{-5} . Se a população pode entrar em equilíbrio, que frequências alélicas podem ser previstas?

SOLUÇÃO

Aqui a mutação e a seleção estão trabalhando em sentidos opostos, logo é previsto um equilíbrio. Tal equilíbrio é descrito pela fórmula

$$\hat{q} = \sqrt{\frac{\mu}{s}}$$

Na presente questão,

$$\mu = 5 \times 10^{-5} \text{ e } s = 1 - W = 1 - 0,9 = 0,1.$$

Logo

$$\hat{q} = \sqrt{\frac{5 \times 10^{-5}}{10^{-1}}} = 0,022$$

$$\hat{p} = 1 - 0,022 = 0,978$$

► Problemas

PROBLEMAS BÁSICOS

- Quais são as forças que podem mudar a frequência de um alelo em uma população?
- Em uma população de camundongos, existem dois alelos do locus A (A_1 e A_2). Os testes mostraram que, nessa população, existem 384 camundongos com o genótipo A_1/A_1 , 210 de A_1/A_2 , e 260 de A_2/A_2 . Quais são as frequências dos dois alelos na população?
- Em uma população de laboratório de *Drosophila* com reprodução aleatória, 4% das moscas têm corpo preto (codificado pelo autossômico recessivo b) e 96% têm corpo marrom (o tipo selvagem, codificado por B). Se essa população for suposta estando em equilíbrio de Hardy-Weinberg, quais são as frequências alélicas de B e b e as frequências genotípicas de B/B e B/b ?
- Em uma população tipo selvagem de besouros da espécie X , você notou que há uma proporção de 3:1 de cobertura de asas brilhantes e foscas. Essa proporção prova que o alelo *shiny* é dominante? (Suponha que os dois estados são causados por dois alelos de um gene.) Caso não, o que isso prova? Como você elucida a situação?
- A adaptabilidade dos três genótipos são $W_{A/A} = 0,9$, $W_{A/a} = 1,0$ e $W_{a/a} = 0,7$.
 - Se a população começou com a frequência alélica $p = 0,5$, qual o valor de p na geração seguinte?
 - Qual o equilíbrio previsto para a frequência alélica?
- Os indivíduos A/A e A/a têm fertilidade igual. Se 0,1% da população é a/a , que pressão de seleção existe contra a/a se a taxa de mutação $A \rightarrow a$ é 10^{-5} ?
- Em um levantamento de tribos nativas norte-americanas no Arizona e Novo México, os albinos eram totalmente ausentes ou muito raros na maioria das tribos (há 1 albino em cada 20.000 norte-americanos brancos).

Entretanto, em três populações nativas americanas, as frequências de albinos são excepcionalmente altas: 1 em cada 277 americanos nativos no Arizona; 1 em cada 140 Jemez no Novo México; e 1 em cada 247 Zuni no Novo México. Todas essas três populações são culturalmente mas não linguisticamente relacionadas. Que fatores possíveis podem explicar a alta incidência de albinos nessas três tribos?

PROBLEMAS DESAFIADORES

- Em uma população, a taxa de mutação de $D \rightarrow d$ é 4×10^{-6} . Se $p = 0,8$ hoje em dia, quanto será p após 50.000 gerações?
- Você está estudando polimorfismo de proteínas em uma população natural de algumas espécies de um organismo haplóide de reprodução sexual. Você isola muitas linhagens de várias partes da área de teste e submete os extratos de cada linhagem a gel de eletroforese. Você cora os géis com um reagente específico para a enzima X e descobre que, na população, há um total de cinco variantes eletroforéticas da enzima X . Você especula que essas variantes representam vários alelos do gene estrutural para a enzima X .
 - Como você poderia demonstrar que sua especulação está correta, tanto genética quanto bioquimicamente? (Você pode fazer cruzamentos, fazer diplóides, correr géis, testar atividades enzimáticas, testar seqüências de aminoácidos, e assim por diante.) Destaque as etapas e conclusões com precisão.
 - Cite pelo menos um outro modo possível de gerar as diferentes variantes eletroforéticas, e explique como você distinguiria essa possibilidade de sua especulação.
- Um estudo feito em 1958 na cidade de mineração de Ashibetsu em Hokkaido, Japão, revelou as frequências

de genótipos de tipo sanguíneo MN (para indivíduos e para casais) mostrados no quadro seguinte:

Genótipo	Nº de indivíduos ou casais
Indivíduos	
L^M/L^M	406
L^M/L^N	744
L^N/L^N	332
Total	1.482
Casais	
$L^M/L^M \times L^M/L^M$	58
$L^M/L^M \times L^M/L^N$	202
$L^M/L^N \times L^M/L^N$	190
$L^M/L^M \times L^N/L^N$	88
$L^M/L^N \times L^N/L^N$	162
$L^N/L^N \times L^N/L^N$	41
Total	741

- a. Mostre se a população está em equilíbrio de Hardy-Weinberg com relação aos tipos sanguíneos MN.
- b. Mostre se a reprodução é aleatória com relação aos tipos sanguíneos MN.
- (O Problema 10 é de J. Kuspira and G.W. Walker, *Genetics: Questions and Problems*. Copyright 1973 by McGraw-Hill.)
11. Considere as populações que têm os genótipos mostrados no quadro seguinte:

População	A/A	A/a	a/a
1	1,0	0,0	0,0
2	0,0	1,0	0,0
3	0,0	0,0	1,0
4	0,50	0,25	0,25
5	0,25	0,25	0,50
6	0,25	0,50	0,25
7	0,33	0,33	0,33
8	0,04	0,32	0,64
9	0,64	0,32	0,04
10	0,986049	0,013902	0,000049

- a. Quais as populações que estão em equilíbrio de Hardy-Weinberg?
- b. Quais são p e q em cada população?
- c. Na população 10, a taxa de mutação de $A \rightarrow a$ foi descoberta como sendo de 5×10^{-6} e a mutação reversa é desprezível. Qual deve ser a adaptabilidade do fenótipo a/a ?
- d. Na população 6, o alelo a é deletério. Além disso, o alelo A é incompletamente dominante; logo, A/A está perfeitamente adaptado, A/a tem uma adaptabilidade de 0,8 e a/a tem uma adaptabilidade de 0,6. Se não houver mutação, quais serão p e q na geração seguinte?
12. O daltonismo resulta de um alelo recessivo ligado ao sexo. Um em cada 10 homens é daltônico.
- a. Que proporção de todas as mulheres é daltônica?

- b. Por que fator o daltonismo é mais comum em homens (ou, quantos homens daltônicos existem para cada mulher daltônica)?
- c. Em que proporção de casamentos o daltonismo afeta metade dos filhos de cada sexo?
- d. Em que proporção de casamentos todas as crianças seriam normais?
- e. Em uma população que não está em equilíbrio, a frequência do alelo para o daltonismo é 0,2 em mulheres e 0,6 em homens. Após uma geração de casamentos aleatórios, que proporção da prole feminina será daltônica? E da prole masculina?
- f. Quais seriam as frequências alélicas na prole de homens e mulheres na parte e?
- (Problema 12 por cortesia de Clayton Person.)
13. Parece claro que a maioria das novas mutações são deletérias. Por quê?
14. A maioria das mutações são recessivas em relação ao tipo selvagem. Dessas mutações raras que são dominantes em *Drosophila*, por exemplo, a maioria é de mutações cromossômicas ou inseparáveis de mutações cromossômicas. Explique por que o tipo selvagem é geralmente dominante.
15. Dez por cento de homens de uma grande população de reprodução aleatória são daltônicos. Um grupo representativo de 1.000 pessoas dessa população migra para uma ilha do Pacífico Sul, onde já existem 1.000 habitantes dos quais 30% dos homens são daltônicos. Supondo que o equilíbrio de Hardy-Weinberg se aplica (nas duas populações originais antes da migração e na população mista imediatamente após a migração), que fração de homens e mulheres pode ser esperada como sendo daltônica na geração imediatamente após a chegada dos migrantes?
16. Usando diagramas de heredogramas, descubra a probabilidade de homozigose por descendência da prole de (a) reproduções genitor-prole; (b) reproduções de primos em primeiro grau; (c) reproduções tia-sobrinho ou tio-sobrinha.
17. Em uma população animal, 20% dos indivíduos são A/A , 60% são A/a e 20% são a/a .
- a. Quais as frequências alélicas nessa população?
- b. Nessa população, o casamento é sempre com *fenótipos similares*, mas é aleatório dentre os fenótipos. Que frequências genotípicas e alélicas prevalecerão na geração seguinte?
- c. Outro tipo de reprodução preferencial ocorre apenas entre fenótipos *diferentes*. Responda a pergunta anterior levando em conta essa restrição.
- d. Qual será o resultado final após muitas gerações de reprodução de cada tipo?
18. Um estoque de *Drosophila* isolado da natureza tem uma média de 36 cerdas abdominais. Por reprodução seletiva apenas dessas moscas com o maior número de cerdas, a média é elevada para 56 cerdas em 20 gerações.
- a. Qual a fonte dessa flexibilidade genética?
- b. O estoque de 56 cerdas é estéril, e, portanto, a seleção é relaxada por várias gerações e o número de cerdas cai para cerca de 45. Por que ele não cai para 36?

- c. Quando a seleção é reaplicada, 56 cerdas logo são atingidas, mas dessa vez o estoque *não* é fértil. Como surgiu essa situação?
19. O alelo *B* é um autossômico dominante deletério. A frequência de indivíduos afetados é $4,0 \times 10^{-6}$. A capacidade reprodutiva desses indivíduos é de cerca de 30% da de indivíduos normais. Avalie μ , a taxa na qual *b* muta para seu alelo deletério *B*.
20. De 31 crianças nascidas de reproduções pai-filha, 6 morreram na lactância, 12 eram muito anormais e morreram na infância, e 13 eram normais. A partir dessa informação, calcule aproximadamente quantos genes letais recessivos nós temos, em média, em nossos genomas humanos. (**Pista:** Se a resposta fosse 1, então a filha teria uma chance de 50% de levar o alelo letal, e a probabilidade de a união produzir uma combinação letal seria de $1/2 \times 1/4 = 1/8$. Logo, 1 não é a resposta.) Considere também a possibilidade de fatalidades não-detectadas no útero em tais reproduções. Como elas afetariam o resultado?
21. Se definirmos o *custo total da seleção* para uma população de genes recessivos deletérios como a perda de adaptabilidade por indivíduo afetado (*s*) multiplicado pela frequência de indivíduos afetados (q^2), então
- $$\text{custo da seleção} = sq^2$$
- a. Suponha que uma população está em equilíbrio entre a mutação e a seleção para um alelo deletério recessivo, onde $s = 0,5$ e $\mu = 10^{-5}$. Qual a frequência de equilíbrio do alelo? Qual o custo da seleção?
- b. Suponha que começamos irradiando membros de uma população de modo que a taxa de mutação dobre. Qual a nova frequência de equilíbrio do alelo? Qual o novo custo da seleção?
- c. Se não mudarmos a taxa de mutação, mas diminuirmos a intensidade de seleção para 0,3, o que ocorrerá com a frequência de equilíbrio e o custo da seleção?

EXPLORANDO GENOMAS Genética Interativa – Megamanual de Tutorial em CD-ROM

Genética de Populações

Essa atividade no CD-ROM de Genética Interativa inclui cinco problemas interativos destinados a melhorar sua com-

preensão do que pode ser aprendido observando a distribuição de genes nas populações.

Genética Quantitativa

18



Flores compostas de *Gaillardia pulchella* mostrando variação quantitativa na cor, no diâmetro e no número de partes das flores. [J. Heywood, *J. Hered.*, May/June 1986.]

Ao final, a meta da Genética é a análise dos genótipos dos organismos. Mas um genótipo pode ser identificado e, portanto, estudado apenas por seu efeito no fenótipo. Reconhecemos dois genótipos como diferentes um do outro porque os fenótipos de seus portadores são diferentes. Os experimentos genéticos básicos, então, dependem da existência de uma relação simples entre o genótipo e o fenótipo. O motivo de os estudos das seqüências de DNA serem tão importantes é que podemos ler diretamente o genótipo.

Em geral, esperamos encontrar um único fenótipo distinguível de cada genótipo, e apenas um único genótipo para cada fenótipo. Na pior das hipóteses, quando um alelo é completamente dominante, pode ser necessário fazer um cruzamento genético simples para distinguir o heterozigoto do homozigoto. Quando possível, os geneticistas evitam estudar genes que tenham apenas penetrância parcial e expressividade incompleta (veja o Capítulo 6) devido à dificuldade em fazer deduções genéticas de tais características. Imagine como teria sido difícil (se não impossível) para Seymour

Perguntas Fundamentais

- Para determinada característica, como respondemos à pergunta: A variação observada na característica é influenciada pela variação genética? São os alelos que se segregam na população que produzem algum efeito diferencial na característica ou toda a variação resulta simplesmente de variação ambiental e ruído de desenvolvimento (veja o Capítulo 1)?
- Se há variação genética, quais são os fenótipos dos vários genótipos em ambientes diferentes?
- Para determinada característica, qual a importância da variação genética como fonte da variação fenotípica total? As normas de reação e os ambientes são tais que toda a variação é uma consequência da diferença ambiental e das instabilidades desenvolvimentais, ou a variação genética predomina?
- Muitos loci (ou apenas alguns) variam com relação a determinada característica? Como eles são distribuídos pelo genoma?

Tópicos

- 18.1 Genes e características quantitativas
- 18.2 Algumas noções estatísticas básicas
- 18.3 Genótipos e distribuição fenotípica
- 18.4 Norma de reação e distribuição fenotípica
- 18.5 Determinando as normas de reação
- 18.6 A herdabilidade de uma característica quantitativa
- 18.7 Quantificando a herdabilidade
- 18.8 Localizando os genes

As características quantitativas variam em uma gama



FIG. 18.1 Herança quantitativa da bráctea da castileja (*Castilleja hispida*). A foto da esquerda mostra os extremos da gama de cores, e a da direita, exemplos da gama de fenótipos.

Benzer estudar mutações dentro do gene *rII* em fagos se o único efeito dos mutantes *rII* fosse uma redução de 5% do tipo selvagem em sua habilidade de crescer em *E. coli* K. Na maior parte, então, o estudo da Genética apresentado nos capítulos precedentes foi o estudo de substituições alélicas que causam diferenças *qualitativas* em diferenças fenotípicas nítidas, tais como flores púrpura *versus* flores brancas.

Entretanto, a maioria das variações entre os organismos é quantitativa, e não qualitativa. O trigo em um campo cultivado ou o mato à beira de uma estrada não são nitidamente classificados em categorias “alta” e “baixa”, assim como os humanos não são distribuídos em categorias de “negros” e “brancos”. A altura, peso, forma, cor, atividade metabólica, taxa reprodutiva e comportamento são características que variam mais ou menos continuamente em uma ampla gama (Figura 18.1). Mesmo quando a característica é intrinsecamente contável (tal como facetas dos olhos ou número de cerdas em *Drosophila*), o número de classes distinguíveis pode ser tão grande que a variação é quase contínua. Se considerarmos indivíduos extremos, digamos, uma planta de milho com 240 cm de altura e outra com 90 cm de altura, um cruzamento entre elas não produzirá um resultado mendeliano. Tal cruzamento de milho produzirá plantas com cerca de 180 cm de altura, com uma clara variação entre irmãos. A F_2 de autofecundação da F_1 não cairá em duas ou três classes distintas de altura nas proporções de 3:1 ou 1:2:1. Em vez disso, a F_2 será continuamente distribuída em altura de um genitor extremo ao outro.

Como estudamos características quantitativas quando elas apresentam uma relação complexa entre o genótipo e o fenótipo? Neste capítulo, veremos que a análise de uma característica de variação contínua pode ser feita por vários tipos de investigação.

► 18.1 Genes e Características Quantitativas

Um exemplo clássico do resultado de cruzamentos entre linhagens que diferem em uma característica quantitativa é o experimento mostrado na Figura 18.2. O tamanho da corola (tubo da flor) foi medido em várias plantas individuais de linhagens puras de *Nicotiana longiflora*, parente do tabaco. A distribuição dos tamanhos da corola de duas linhagens parentais é mostrada no topo do painel da Figura 18.2. A diferença entre as duas linhagens é genética, mas a variação entre plantas individuais dentro de cada linhagem é o resultado de variação ambiental descontrolada e ruído desenvolvimental. As plantas F_1 , cujo tamanho médio da corola é muito próximo entre duas linhagens parentais, também varia de uma para outra devido à variação ambiental e desenvolvimental. Na F_2 , o tamanho médio da corola permanece essencialmente inalterado em comparação ao da F_1 , mas há um grande aumento na variação porque agora há segregação das diferenças genéticas que foram introduzidas a partir das duas linhagens parentais. Uma demonstração de que pelo menos parte dessa variação é o resultado de diferenças genéticas entre as plantas F_2 é vista na F_3 . Foram escolhidos pares diferentes de genitores das quatro partes diferentes da distribuição de F_2 e cruzados para produzir a geração seguinte, F_3 . Em cada caso, a média de F_3 é próxima do valor dessa parte da distribuição F_2 da qual foram colhidos seus genitores.

O resultado do cruzamento é claramente diferente dos resultados obtidos quando é feito um cruzamento entre indivíduos que diferem em seu estado alélico para um gene com efeitos fenotípicos nítidos. A prole não se distribui em claras proporções mendelianas de 1:2:1, e há muito mais variação individual dentro de cada geração da prole. Mendel obteve seus resultados simples porque trabalhou com variação

A variação em características quantitativas aumenta na prole

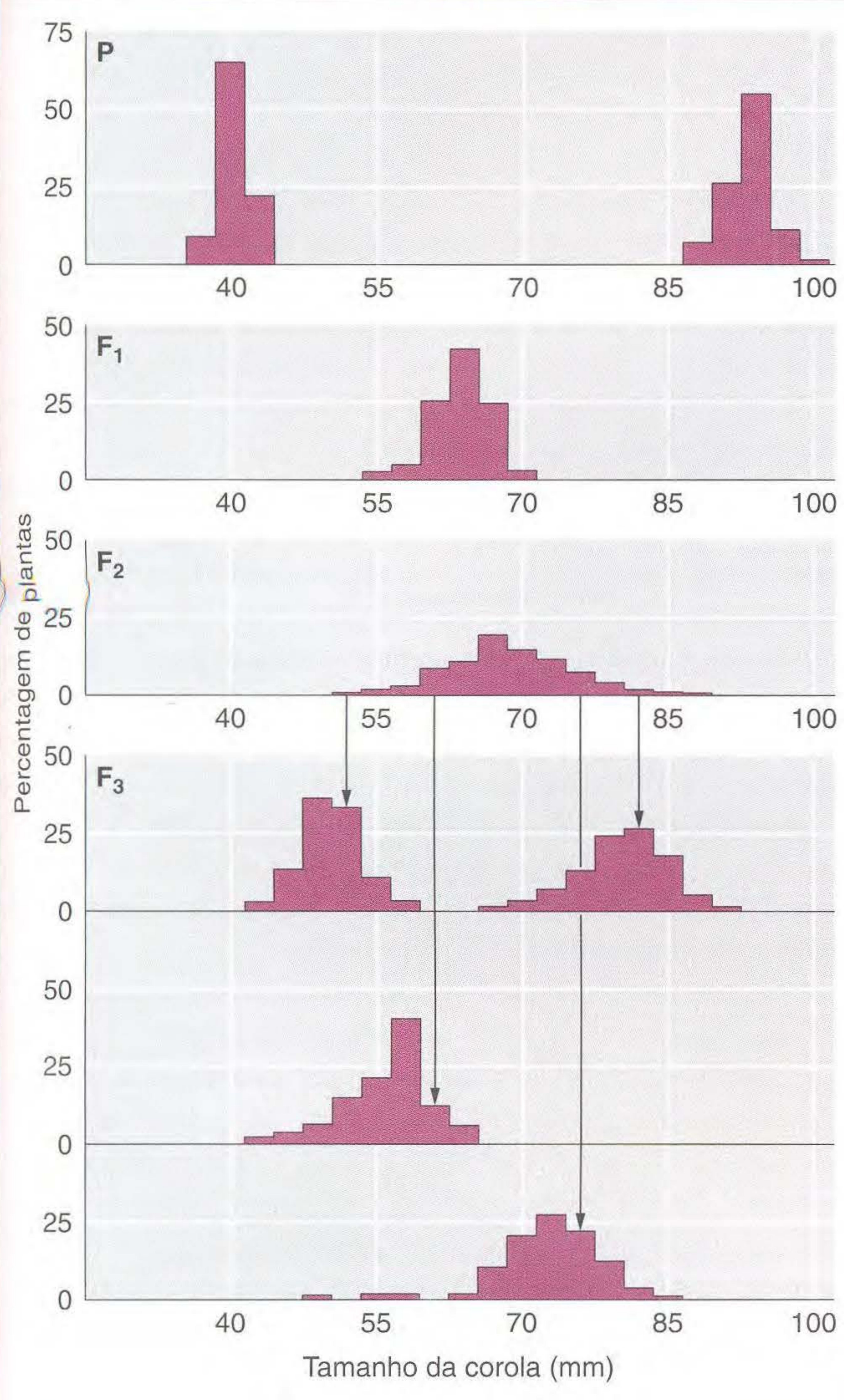


FIG. 18.2 Resultados dos cruzamentos entre linhagens de *Nicotiana longiflora* que diferem no tamanho da corola. Os gráficos mostram (de cima para baixo) a distribuição da frequência dos tamanhos de corola nas duas linhagens parentais (P); a distribuição da frequência dos tamanhos da corola na F_1 ; a distribuição da frequência na F_2 ; e a distribuição da frequência em quatro cruzamentos de F_3 feitos tomando genitores das quatro partes indicadas na distribuição de F_2 . [De K. Mather, *Biometrical Genetics*. Methuen, 1959. Dados de E. M. East, *Genetics* 1, 1916, 164-176.]

des hortícolas de ervilhas de jardim que diferem uma da outra por diferenças alélicas únicas que têm marcantes efeitos nos fenótipos. O comportamento dos cruzamentos vistos na Figura 18.2 não é exceção, mas sim a regra para a maioria das características na maior parte das espécies. Se Mendel tivesse feito seus experimentos na variação natural de pragas em seu jardim, e não em variedades anormais de ervilhas, ele nunca teria descoberto nenhuma de suas leis da hereditariedade. Em geral, o tamanho, forma, cor, atividade fisiológica e comportamento não se distribuem de modo simples nos cruzamentos.

O fato de que a maioria das características variam continuamente não significa que sua variação seja o resultado de alguns mecanismos genéticos diferentes dos que se aplicam aos genes mendelianos que estudamos nos capítulos anteriores. A continuidade do fenótipo é resultado de dois fenômenos. Primeiro, cada genótipo não tem uma única expressão fenotípica, mas sim uma que depende do ambiente no qual o organismo se desenvolve, e o fenótipo também varia devido a eventos de desenvolvimento aleatórios. Como resultado, as diferenças fenotípicas entre as classes genotípicas tornam-se indistintas, e não somos capazes de atribuir determinado fenótipo claramente a determinado genótipo.

Segundo, deve haver muitos loci segregantes tendo alelos para fazer uma diferença no fenótipo sob observação. Suponha, por exemplo, que cinco loci igualmente importantes afetam o número de flores que irão desenvolver-se em uma planta anual, e que cada locus tem dois alelos (chamemos de + e -). Por simplificação, suponhamos também que não há dominância e que um alelo + adiciona uma flor, enquanto um alelo - não adiciona nada. Assim, há $3^5 = 243$ genótipos diferentes possíveis [três genótipos possíveis (+/+, +/- e -/-) em cada um dos cinco loci], variando de

$$\frac{+ + + + +}{+ + + + +} \text{ até } \frac{- - - - -}{- - - - -}$$

mas existem apenas 11 classes fenotípicas (10, 9, 8, ..., 0) porque muitos dos genótipos terão os mesmos números de alelos + e -. Embora só exista um genótipo com 10 alelos + e, portanto, um valor fenotípico médio de 10, existem 51 genótipos diferentes com 5 alelos + e 5 -; por exemplo,

$$\frac{+ + + + -}{+ - - - -} \text{ e } \frac{+ + - + -}{+ + - - -}$$

Assim, muitos genótipos diferentes podem ter o mesmo fenótipo médio. Por outro lado, devido à variação ambiental, dois indivíduos do mesmo genótipo podem não ter o mesmo fenótipo. Essa falta de correspondência entre genótipo e fenótipo obscurece o mecanismo mendeliano subjacente.

Se não pudermos estudar o comportamento de fatores mendelianos que controlam diretamente tais características, então o que podemos aprender sobre sua genética? Claramente, os métodos usados para analisar características qualitativas, tais como examinar as proporções da prole em um cruzamento genético, não funcionarão para características quantitativas. Em vez disso, temos que usar métodos estatísticos para fazer previsões sobre a herança de fenótipos na ausência de conhecimentos sobre os genótipos subjacentes. Esse enfoque é conhecido como genética quantitativa. O estudo da genética de características de varia-

ção contínua, **genética quantitativa**, implica responder às seguintes questões:

1. A variação observada em uma característica é influenciada pela variação genética? Toda a variação é simplesmente o resultado de variação ambiental e ruído de desenvolvimento (veja o Capítulo 1)? Ou existem três alelos segregando na população que produzem algum efeito diferencial na característica?
2. Se houver variação genética, quais são os fenótipos dos vários genótipos em ambientes diferentes?
3. Qual a importância da variação genética como fonte de variação fenotípica total? Quase toda a variação é uma consequência de diferenças ambientais e instabilidades desenvolvimentais ou a variação genética predomina?
4. Muitos loci (ou apenas alguns) contribuem para a variação na característica? Como eles estão distribuídos pelo genoma? Qual a base molecular e sua influência no fenótipo?

No final, o propósito de responder a essas perguntas é ser capaz de prever que tipos de prole serão produzidos pelos cruzamentos de fenótipos diferentes.

A precisão com a qual essas perguntas podem ser feitas e respondidas varia muito. Em organismos experimentais, por outro lado, é relativamente simples determinar se existe alguma influência genética, mas experimentos muito trabalhosos são necessários para localizar os genes (mesmo que aproximadamente). Em humanos, por outro lado, é extremamente difícil responder até mesmo a questão da presença de influência genética para a maioria das características, porque é quase impossível separar os efeitos ambientais dos genéticos em um organismo que não pode ser manipulado experimentalmente. Como consequência, sabemos muito sobre a genética do número de cerdas em *Drosophila*, mas praticamente nada sobre a genética de características humanas complexas. Algumas (tais como cor da pele) claramente são influenciadas por genes, enquanto outras (tais como a linguagem específica falada) claramente não. Neste capítulo, desenvolveremos a estatística básica e os conceitos genéticos necessários para responder essas perguntas e dar alguns exemplos das aplicações desses conceitos em características particulares em espécies particulares.

► 18.2 Algumas Noções Estatísticas Básicas

Para considerarmos as respostas a essas perguntas sobre os tipos mais comuns de variação fenotípica, a variação quantitativa, precisamos primeiro examinar várias ferramentas estatísticas que são essenciais ao estudo da variação quantitativa.

Distribuições estatísticas

Para simples variação que depende apenas de diferenças alélicas em um único locus, a prole de um cruzamento cairá em várias classes fenotípicas distintas. Por exemplo, um cruzamento entre uma planta de flor vermelha e uma planta

de flor branca deverá produzir todas as plantas de flores vermelhas ou, se for um retrocruzamento de uma F_1 com o genitor de flores brancas, $\frac{1}{2}$ das plantas de flores vermelhas e $\frac{1}{2}$ das plantas de flores brancas. Entretanto, precisamos de um modo diferente de descrição para características quantitativas. Se as alturas de um grande número de estudantes masculinos de graduação forem medidas até os próximos 5 centímetros (cm), elas irão variar (digamos, entre 145 e 195 cm), mas muito mais alunos cairão nas classes de medidas médias (digamos, 170, 175 e 180 cm) do que nas classes dos dois extremos. Tal descrição de um conjunto de medidas quantitativas é conhecida como uma **distribuição estatística**.

Podemos fazer um gráfico de tais medidas representando cada classe medida como uma barra, cuja altura é proporcional ao número de indivíduos nessa classe, como mostrado na Figura 18.3a. Tal gráfico de número de indivíduos observados *versus* classe de medida é chamado de **histograma de frequência**. Agora, suponha que medimos cinco vezes mais indivíduos, cada um até o centímetro mais próximo, de modo que sejam divididos em classes ainda menores, produzindo um histograma como o mostrado na Figura 18.3b. Se continuamos esse processo, tornando cada medida mais fina mas proporcionalmente aumentando o número de indivíduos medidos, o histograma terá um aspecto contínuo, como na Figura 18.3c. Tal curva contínua é chamada de **função de distribuição** de medida na população.

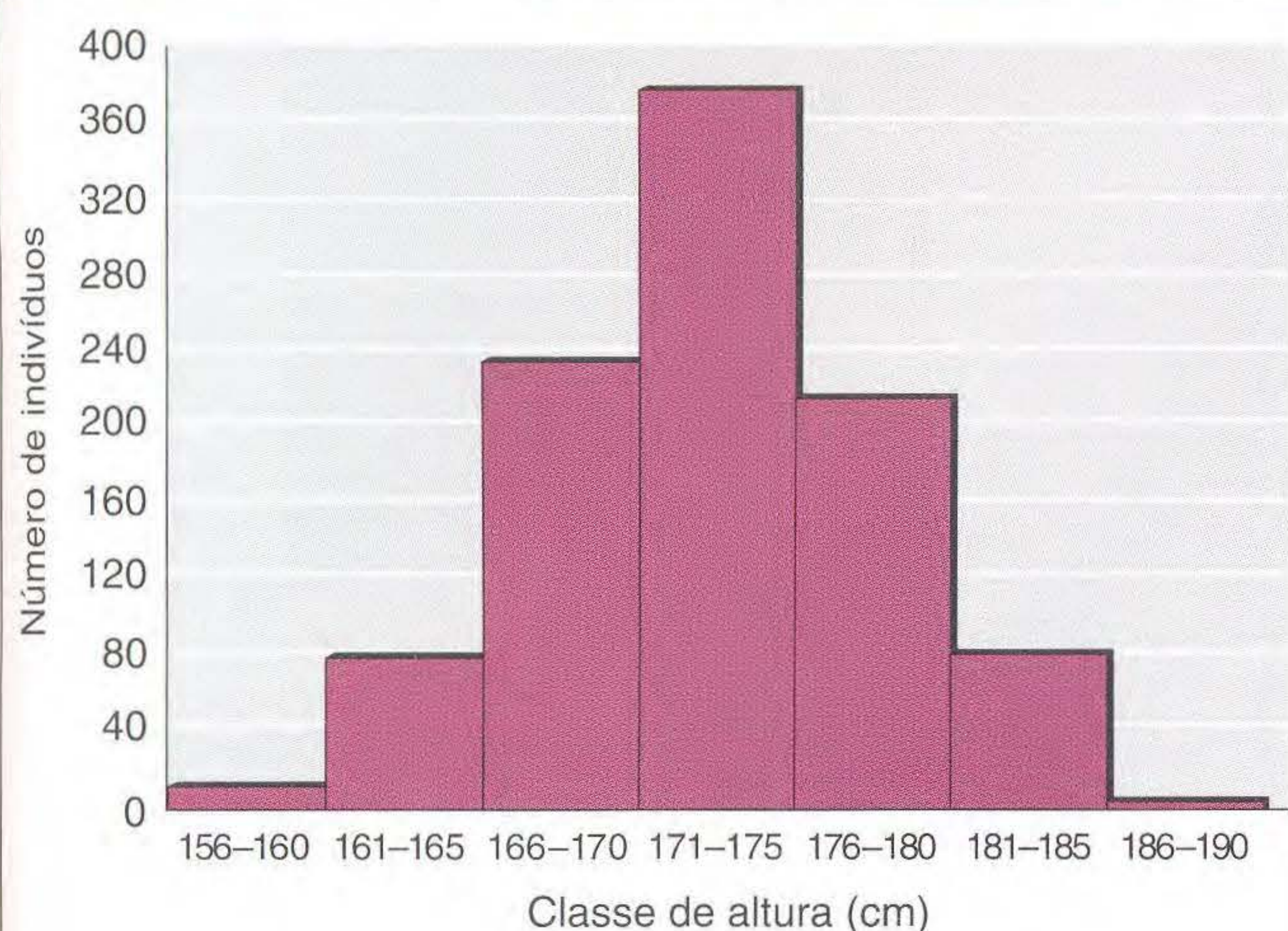
A função de distribuição é uma idealização da distribuição da frequência real de uma medida em qualquer população real, pois nenhuma medida pode ser obtida com precisão infinita ou em um número ilimitado de indivíduos. Além disso, a própria característica pode ser intrinsecamente descontínua, pois é a contagem de algum número de objetos distintos, tais como as facetas dos olhos ou as cerdas. Às vezes é conveniente, entretanto, desenvolver conceitos usando essa curva idealizada como uma abreviação para um histograma mais complexo observado.

Medidas estatísticas

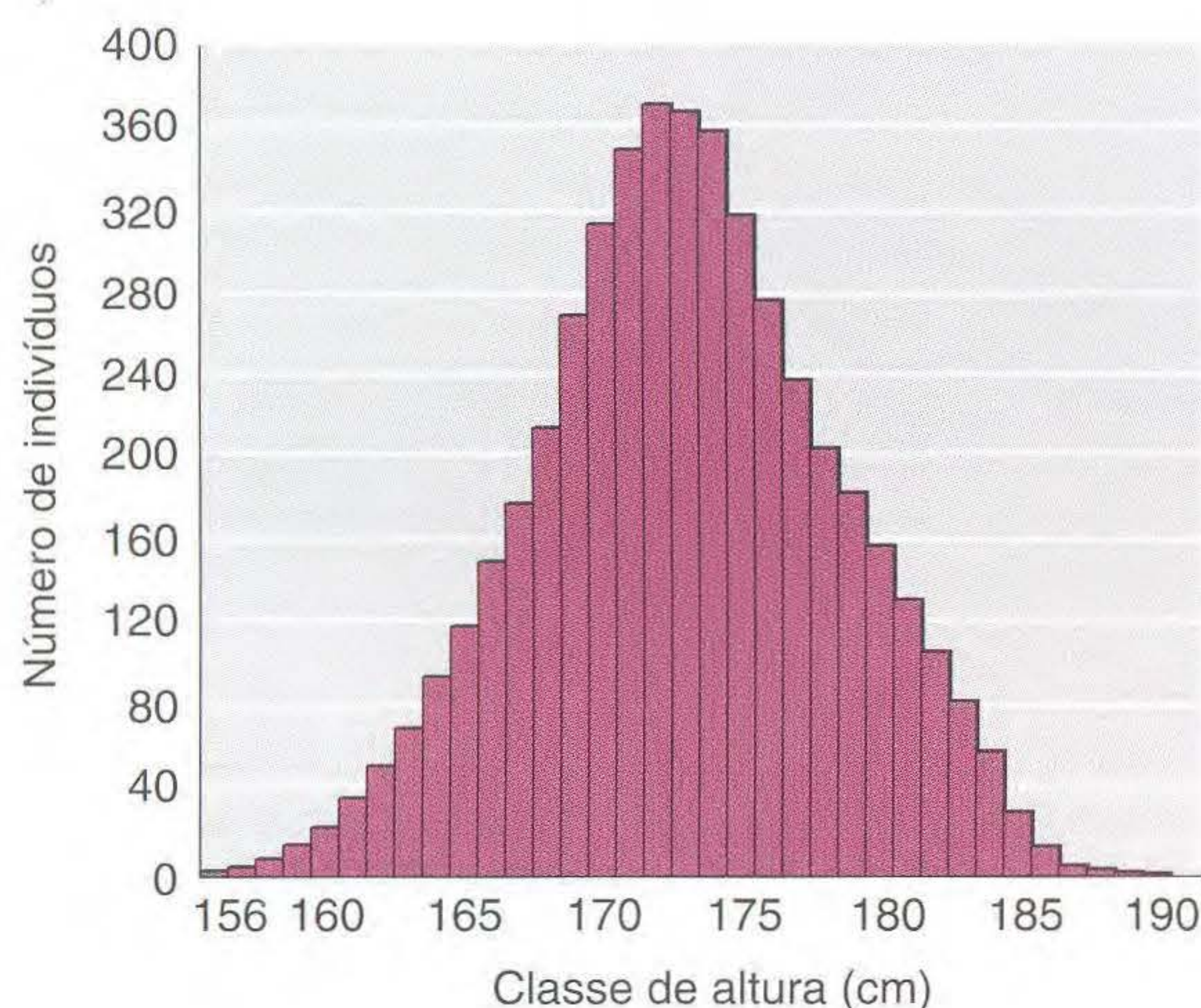
Embora uma distribuição estatística contenha todas as informações que precisamos sobre um conjunto de medidas, em geral é útil fragmentar essa informação em alguns números característicos que levam à informação necessária sobre a distribuição sem dá-las em detalhes. Existem várias perguntas sobre as distribuições de altura de alunos homens, por exemplo, que poderíamos querer responder:

1. Onde a distribuição está situada ao longo da faixa de valores possíveis? Nossos valores observados de altura, por exemplo, estão mais próximos de 100 ou de 200 cm? Essa pergunta pode ser respondida com uma medida da **tendência central**.
2. O quanto de variação existe entre as medidas individuais? Elas estão todas concentradas ao redor da medida central ou variam amplamente em uma grande faixa? Essa pergunta pode ser respondida com a medida da **dispersão**.
3. Se estivermos considerando mais de uma quantidade medida, como os valores das diferentes quantidades estão rela-

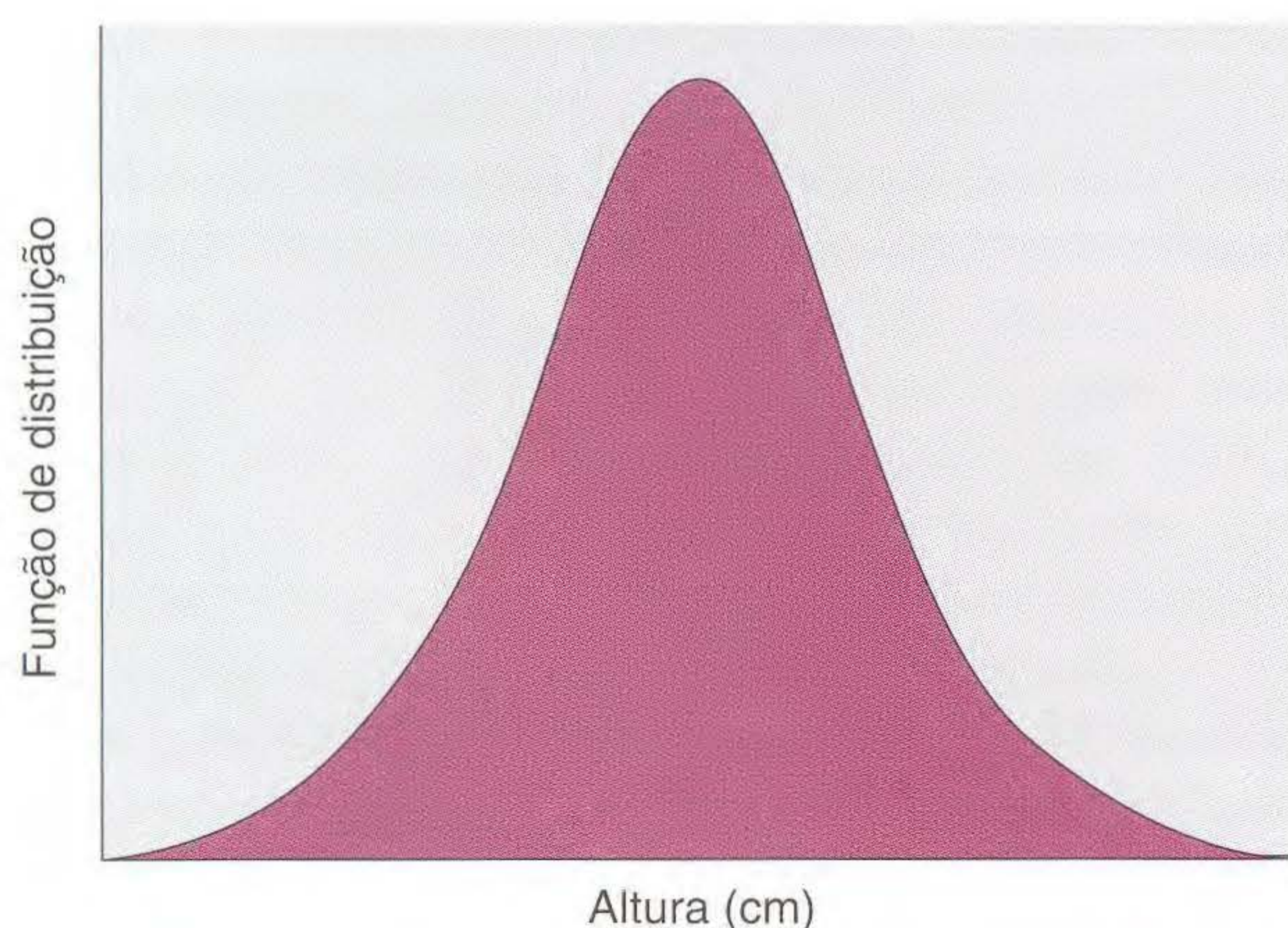
As características quantitativas são descritas pela distribuição da frequência



(a)



(b)



(c)

FIG. 18.3 Distribuições da frequência para altura de alunos de graduação. (a) Um histograma de frequência com intervalos de classes de 5 cm; (b) um histograma de frequência com intervalos de classe de 1 cm; (c) a distribuição contínua.

cionados? Os genitores mais altos, por exemplo, têm filhos mais altos? Caso tenham, veríamos esse achado como uma evidência de que os genes influenciam a altura. Assim, precisamos de medidas de **relação** entre as medidas.

Entre as medidas mais comumente usadas de tendência central estão a **moda**, que é a observação mais frequente, e a **média**, que é a média aritmética das observações. A dispersão de uma distribuição é quase sempre medida pela **variança**, que é o quadrado do desvio médio das observações de sua média, ou o **desvio padrão**, que é a raiz quadrada da variância. A relação entre variáveis diferentes é medida por sua **covariança** e sua **correlação**. A **covariança** entre duas variáveis é o produto médio do desvio de uma variável de sua própria média vezes o desvio da outra variável de sua própria média. A **correlação** é a covariança entre duas variáveis dividida pelo produto dos desvios padrões das duas variáveis. Como medida de relação, a correlação tem a vantagem de variar entre +1 para variáveis positivamente relacionadas, até 0 para variáveis que não têm relação média e para -1 para variáveis que estão perfeitamente relacionadas negativamente.

Essas medidas comuns são discutidas em detalhe no Apêndice Estatístico na análise estatística ao final deste capítulo. A discussão detalhada desses conceitos estatísticos é colocada em uma seção separada para não interromper o fluxo da lógica ao considerarmos a genética quantitativa. Não se deve supor, entretanto, que a compreensão desses conceitos estatísticos seja de algum modo secundária. Uma compreensão apropriada de genética quantitativa requer o entendimento das bases da análise estatística discutidas no Apêndice Estatístico.

► 18.3 Genótipos e Distribuição Fenotípica

A diferença crítica entre características quantitativas e mendelianas

Usando os conceitos de distribuição, média e variância, podemos compreender a diferença entre características genéticas quantitativas e mendelianas.

Suponha que uma população de plantas contém três genótipos, cada um dos quais tem algum efeito diferencial na taxa de crescimento. Além disso, suponha que há alguma variação ambiental de planta a planta porque o solo no qual a população está crescendo não é homogêneo, e existe algum ruído desenvolvimental (veja o Capítulo 1). Para cada genótipo, haverá uma distribuição separada de fenótipos com uma média e uma variância que dependem do genótipo e do conjunto de ambientes. Suponha que essas distribuições se pareçam com as distribuições de três alturas na Figura 18.4a. As três distribuições estão concentradas em três lugares diferentes na escala de altura da planta, indicando diferenças na altura média. As três distribuições também têm quantidades diferentes de dispersão, o que resulta em ter variâncias diferentes. Finalmente, suponha que a população consiste em uma mistura de três genótipos, mas em proporções desiguais 1:2:3 ($a/a:A/a:A/A$).

As distribuições fenotípicas somam para produzir uma distribuição geral com uma única moda

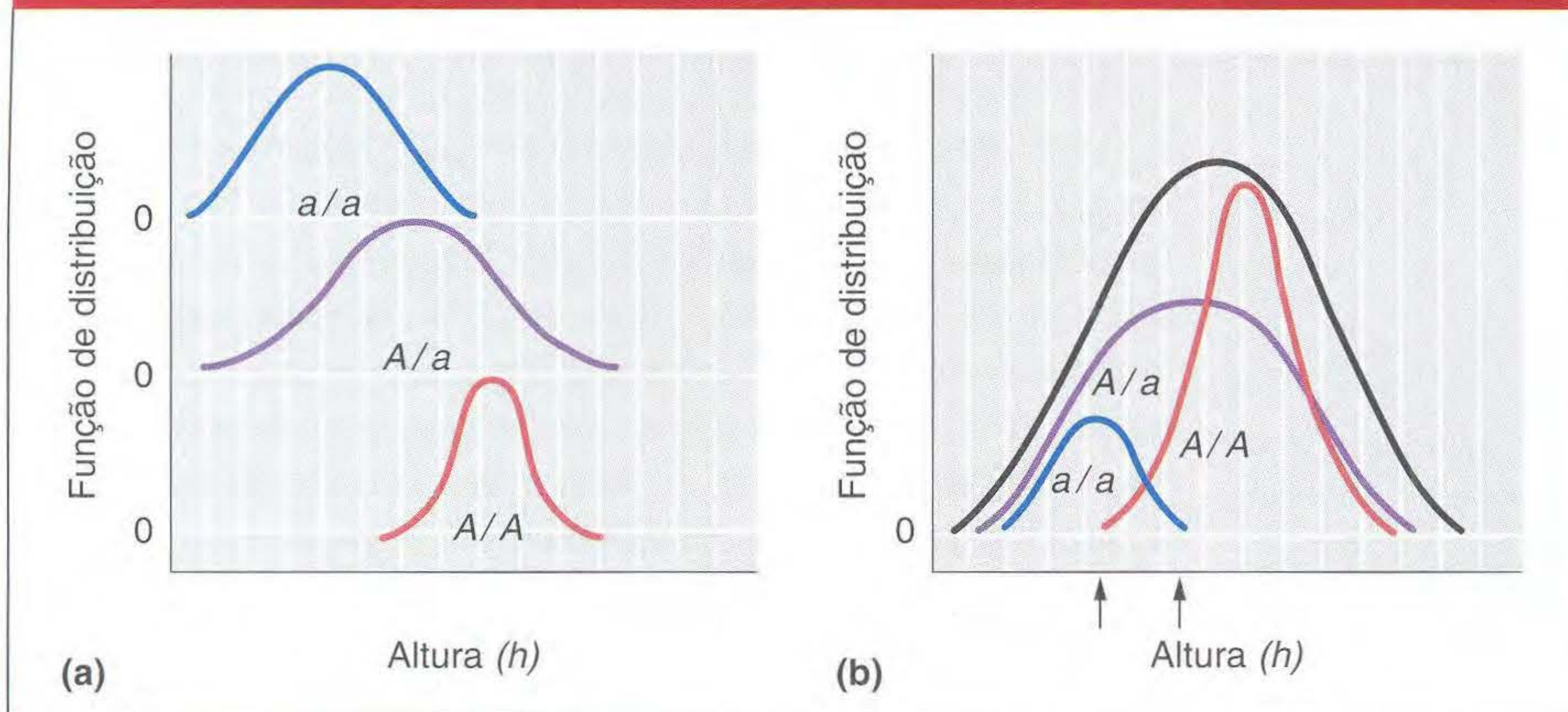


FIG. 18.4 (a) Distribuições fenotípicas de genótipos de três plantas. (b) Uma distribuição fenotípica da população total (linha preta) pode ser obtida somando as três distribuições genotípicas em uma proporção 1:2:3 ($a/a : A/a : A/A$).

Sob essas circunstâncias, a distribuição fenotípica de plantas individuais na população como um todo se parecerá com a linha preta na Figura 18.4b, que é o resultado da soma de três distribuições genotípicas separadas, avaliadas por suas frequências na população. Essa avaliação por frequência é indicada na Figura 18.4b pelas alturas diferentes das distribuições componentes. A média dessa distribuição total é a média das três médias genotípicas, novamente avaliadas pelas frequências dos genótipos na população. A variância da distribuição total é produzida, em parte, pela variação ambiental dentro de cada genótipo e, em parte, pelas médias ligeiramente diferentes dos três genótipos.

Duas características da distribuição total são notáveis. Primeiro, há apenas uma única moda, a observação mais frequente representada pela localização do eixo do pico da curva. Apesar da existência de três distribuições genotípicas separadas subjacentes, a distribuição da população como um todo não revela as modas separadas. Segundo, qualquer planta individual cuja altura esteja entre as duas setas pode ter qualquer um dos três genótipos, porque os fenótipos dos três genótipos são muito superpostos. O resultado é que não podemos fazer uma simples análise mendeliana para determinar o genótipo de uma planta individual. Por exemplo, suponha que os três genótipos sejam os dois homozigotos e o heterozigoto para um par de alelos em um locus. Seja a/a o homozigoto baixo e A/A o alto, com o heterozigoto tendo altura intermediária. Como as distribuições fenotípicas se superpõem muito, não podemos saber a qual genótipo determinada planta pertence. Contrariamente, se cruzarmos uma homozigota a/a e uma heterozigota A/a , a prole não cairá em duas classes distintas, A/a e a/a , em uma proporção de 1:1, mas cobrirá quase toda a faixa de fenótipos. Assim, não podemos saber, olhando a prole, que o cruzamento é de fato $a/a \times A/a$, e não $a/a \times A/A$ ou $A/a \times A/a$.

Suponha que cultivemos as plantas hipotéticas da Figura 18.4 em um ambiente que exagere as diferenças entre os genótipos — por exemplo, duplicando a taxa de crescimento de todos os genótipos. Ao mesmo tempo, fomos muito cuidadosos em dar a todas as plantas exatamente o mesmo ambiente. Então, a variância fenotípica de cada genótipo separado

seria reduzida porque todas as plantas foram cultivadas em condições idênticas. Ao mesmo tempo, as diferenças fenotípicas entre os genótipos seriam exageradas pelo crescimento mais rápido (Figura 18.5a). O resultado (Figura 18.5b) seria uma separação da população como um todo em três distribuições fenotípicas não-superpostas, cada uma característica de um genótipo. Podemos agora fazer uma análise mendeliana perfeitamente convencional da altura das plantas. Uma “característica quantitativa” foi convertida em uma “qualitativa”. Essa conversão foi feita encontrando um meio de tornar a diferença entre as médias dos genótipos grande comparada com a variação dentro dos genótipos.

Uma característica quantitativa é aquela cujas diferenças fenotípicas médias entre os genótipos são pequenas se comparadas com a variação entre indivíduos dentro dos genótipos.

Número de genes e características quantitativas

A variação contínua em uma característica às vezes é admitida como sendo necessariamente causada por um grande número de genes segregantes, e, assim, a variação contínua é tida como evidência de que a característica é controlada por muitos genes. Essa **hipótese de fatores múltiplos** (de que um grande número de genes, cada um com um pequeno efeito, estão segregando para produzir variação quantitativa) há muito tem sido o modelo básico da genética quantitativa, mas, como foi mostrado, essa hipótese não é necessariamente verdadeira. Se a diferença entre as médias genotípicas é pequena comparada com a variância ambiental, então mesmo um caso simples de um gene-dois alelos pode resultar em variação fenotípica contínua.

Se uma faixa de uma característica é limitada, e se muitos loci segregantes a influenciam, então esperamos que a característica apresente variação contínua, porque cada substituição alélica deve corresponder apenas a uma pequena diferença na característica. É importante lembrar, entretanto, que o *número* de loci segregantes que influencia uma característica não é o que separa as características quantitativas e

Convertendo uma característica "quantitativa" em uma "qualitativa"

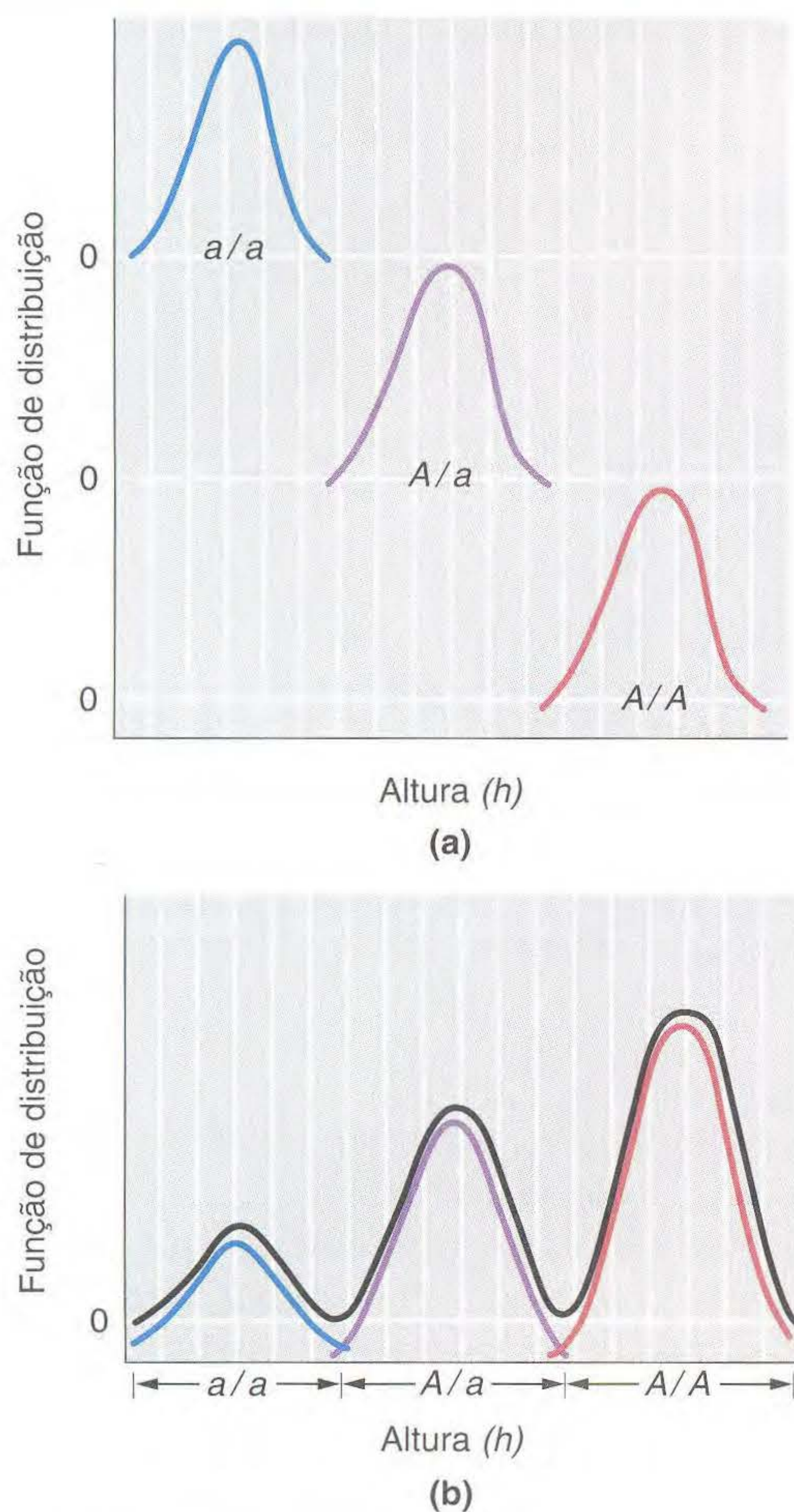


FIG. 18.5 Distribuições fenotípicas dos três genótipos de plantas mostrados na Figura 18.4 quando cultivadas em ambientes cuidadosamente controlados. O resultado é uma variação fenotípica menor com cada genótipo e uma diferença maior entre os genótipos. As alturas das distribuições individuais na parte *b* são proporcionais às frequências de genótipos na população.

qualitativas. Mesmo na ausência de grande variação ambiental, são necessários apenas alguns loci geneticamente variantes para produzir uma variação que é indistinguível do efeito de muitos loci de pequenos efeitos. Como exemplo, podemos considerar um dos primeiros experimentos em genética quantitativa: o de Wilhelm Johannsen em linhagens puras. Por **endogamia** (reprodução entre parentes), Johannsen produziu 19 linhagens homozigotas de plantas de feijão de uma população geneticamente heterogênea. Cada linhagem tinha um peso médio de semente característico. Esses pesos variavam muito, de 0,64 g da linhagem mais pesada até 0,35 g por semente para a mais leve. Suponha que todas essas linhagens fossem geneticamente diferentes. Nesse caso, os resultados de Johannsen seriam incompatíveis com o modelo simples de um locus-dois alelos da ação gênica. Se a população original estivesse segregando para dois alelos A e a , todas as linhagens

endogâmicas derivadas dessa população cairiam em uma das duas classes: A/A ou a/a . Se, em contraste, houvesse, digamos, 100 loci, cada um com um pequeno efeito, segregando na população original, então seria produzido um grande número de linhagens endogâmicas diferentes, cada uma com uma combinação de homozigotos em loci diferentes.

Entretanto, não precisamos de um número grande de loci para obter os resultados observados por Johannsen. Se houvesse apenas cinco loci, cada um com três alelos, então $3^5 = 243$ tipos de homozigotos diferentes poderiam ser produzidos pelo processo de endogamia. Se fizemos 19 linhagens endocruzadas ao acaso, há uma boa chance (cerca de 50%) de que cada uma das 19 linhagens pertençam a uma das 243 classes diferentes. Assim, os resultados experimentais de Johannsen podem ser facilmente explicados por um número relativamente pequeno de genes. Portanto, não existe uma linha divisória entre as características multigênicas e outras características. É seguro dizer que nenhuma característica fenotípica acima do nível da sequência de aminoácidos em um polipeptídeo é influenciada por apenas um gene. Além disso, as características influenciadas por muitos genes não são igualmente influenciadas por todos eles. Alguns genes terão efeitos importantes em uma característica; outros, pequenos efeitos.

A diferença crítica entre características quantitativas e mendelianas não é o número de loci segregantes, mas o tamanho das diferenças fenotípicas comparadas com a variação individual dentro das classes genotípicas.

► 18.4 Norma de Reação e Distribuição Fenotípica

O fenótipo de um organismo depende não só de seu genótipo, mas também do ambiente em que ele vive em vários estágios críticos de seu desenvolvimento. Para determinado genótipo, fenótipos diferentes irão desenvolver-se em ambientes diferentes. A relação entre o ambiente e o fenótipo para determinado genótipo é chamada de **norma de reação** do genótipo. A norma de reação de um genótipo com relação a alguma variável ambiental — digamos, temperatura — pode ser vista em um gráfico mostrando o fenótipo como uma função dessa variável, como exemplificado na Figura 18.6, as normas de reação do número de cerdas abdominais para genótipos diferentes de *Drosophila*.

A distribuição fenotípica de uma característica, como vimos, é uma função das diferenças fenotípicas médias entre genótipos e da variação fenotípica entre indivíduos genotipicamente idênticos. Mas, como as normas de reação na Figura 18.6 mostram, ambas são funções dos ambientes nos quais os organismos se desenvolvem e vivem. Para determinado genótipo, cada ambiente resultará em determinado fenótipo (no momento, ignorando o ruído de desenvolvimento). Assim, para qualquer determinado genótipo, uma *distribuição de ambientes* resultará em uma *distribuição de fenótipos*.

Uma norma de reação é a relação entre o ambiente e o fenótipo

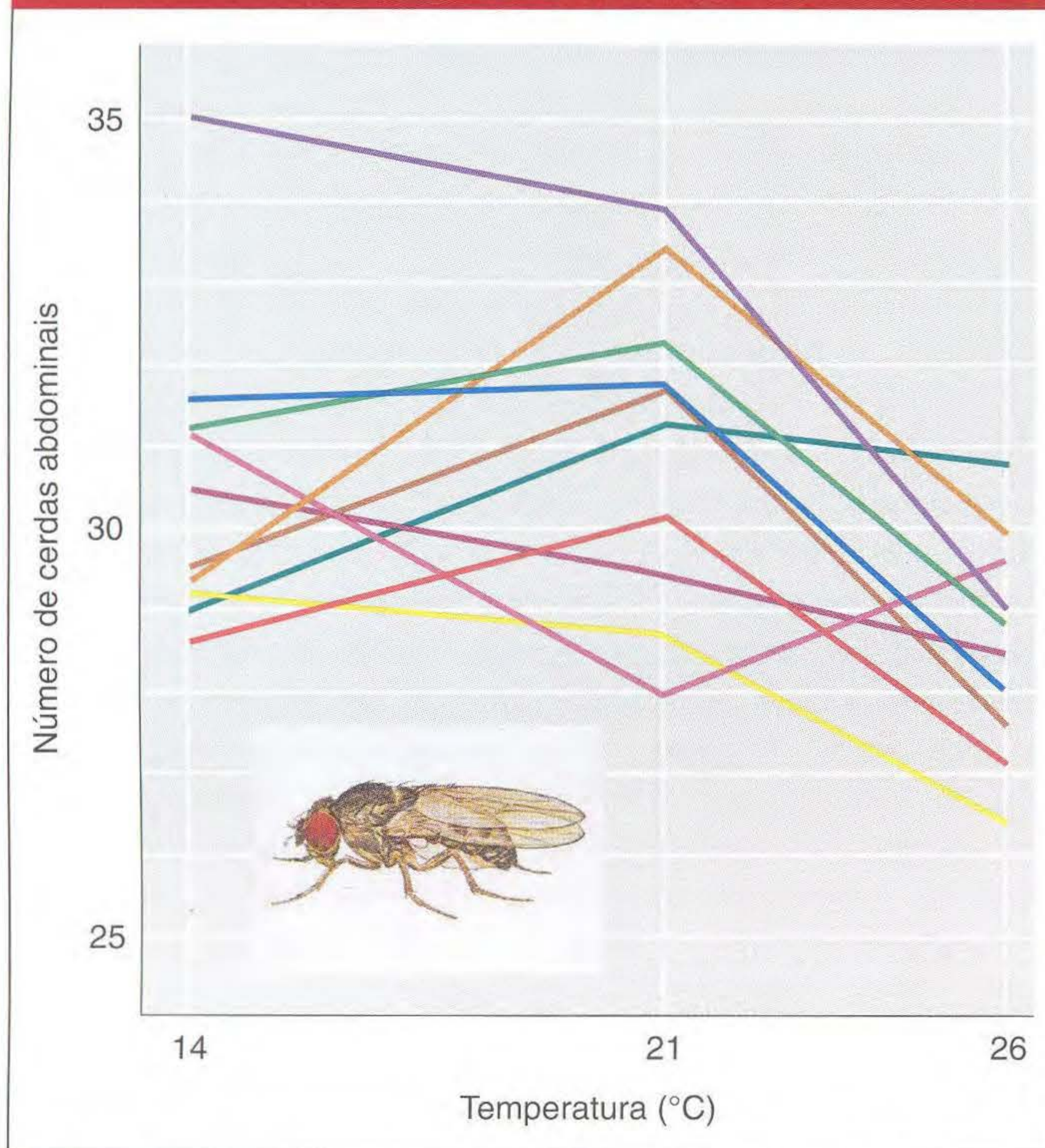


FIG. 18.6 Normas de reação do número de cerdas em *Drosophila*. O número de cerdas abdominais em genótipos homozigotos diferentes de *Drosophila pseudoobscura* em três temperaturas diferentes. Cada linha colorida representa um genótipo diferente. [Dados por cortesia de A.P. Gupta. Image: Plate IV. University of Texas Publication 4313, *Studies in the Genetics of Drosophila III: The Drosophilidae of the Southwest*, de J.T. Patterson. Cortesia de Life Sciences Library, University of Texas, Austin.]

O modo pelo qual ambientes diferentes afetam o fenótipo de um organismo depende da norma de reação, como mostrado na Figura 18.7, no qual o eixo horizontal representa o ambiente (digamos, a temperatura) e o eixo vertical representa o fenótipo (digamos, a altura da planta). A curva da norma de reação para um genótipo mostra como cada determinada temperatura resulta na altura de determinada planta. Essa norma de reação converte a distribuição de ambientes em uma distribuição de fenótipos. Assim, por exemplo, na Figura 18.7, a linha tracejada de 18°C que indica o eixo ambiental horizontal é refletida na curva da norma de reação da altura correspondente de uma planta no eixo vertical do fenótipo, e, assim por diante, para cada temperatura. Se um grande número de plantas se desenvolve a, digamos, 20°C, então um grande número de plantas terá o fenótipo que corresponde a 20°C, como mostrado pela linha tracejada de 20°C. Se apenas um pequeno número se desenvolve a 18°C, algumas plantas terão a altura correspondente da planta. Em outras palavras, a distribuição da frequência dos ambientes desenvolvimentais será refletida como distribuição da frequência de fenótipos como determinado pela forma da curva de norma de reação. É como se um observador, ficando no eixo vertical do fenótipo, estivesse vendo a distribuição ambiental não

Resultados de um estudo de norma de reação

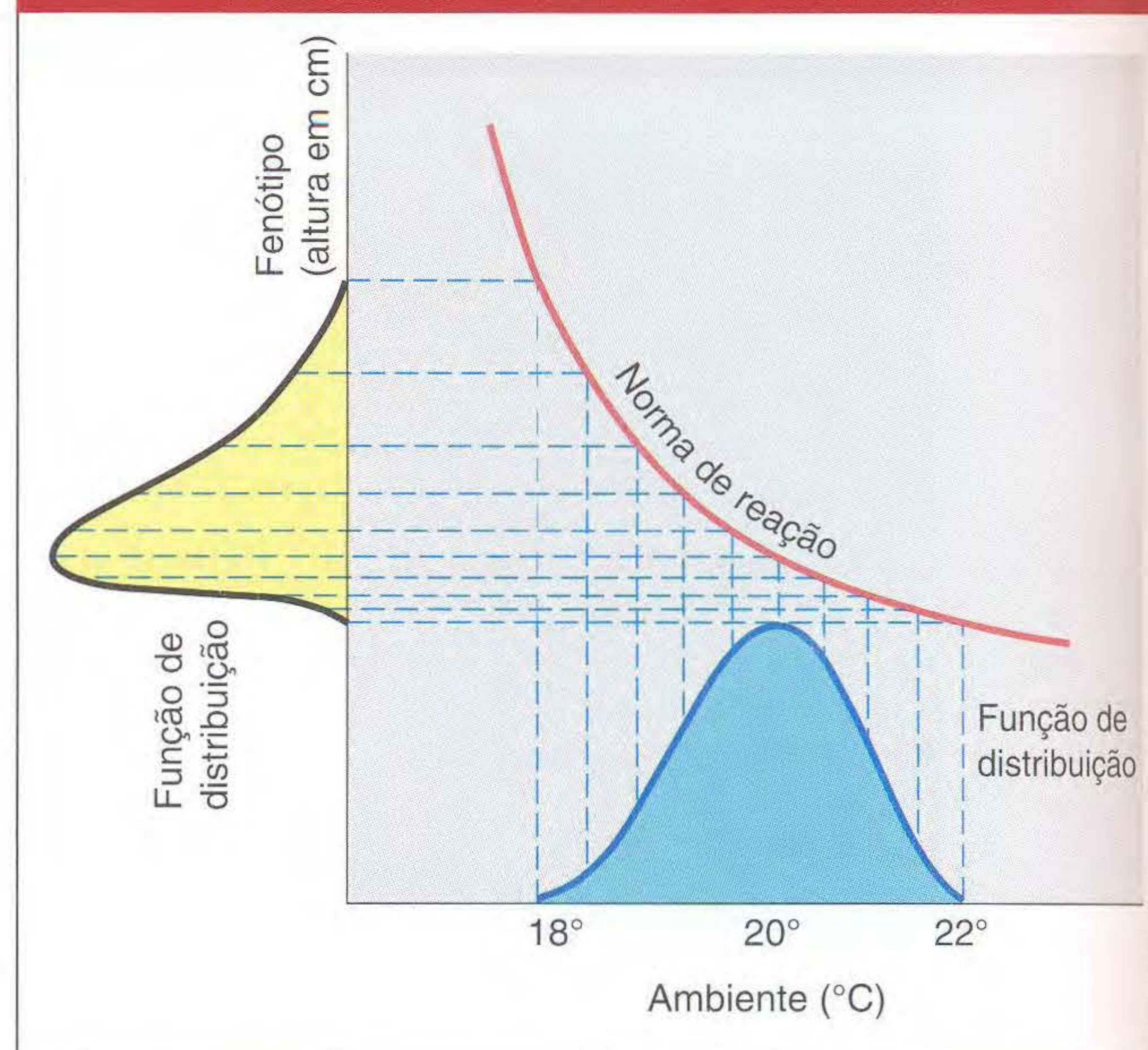


FIG. 18.7 A distribuição de ambientes no eixo horizontal é convertida na distribuição de fenótipos do eixo vertical pela norma de reação de um genótipo.

diretamente, mas refletida em um espelho curvo da norma de reação. A forma de sua curvatura determinará como a distribuição ambiental é distorcida no eixo do fenótipo. Assim, a norma de reação na Figura 18.7 cai muito rapidamente em temperaturas mais baixas (o fenótipo muda acentuadamente com pequenas mudanças na temperatura), mas achata-se em temperaturas altas, mostrando que a altura da planta é muito menos sensível a diferenças de temperatura nas temperaturas mais altas. O resultado é que a distribuição ambiental simétrica é convertida em distribuição fenotípica assimétrica com uma longa cauda nas plantas mais altas, correspondendo a temperaturas mais baixas.

A distribuição dos ambientes é refletida biologicamente como uma distribuição de fenótipos. A transformação da distribuição ambiental em distribuição fenotípica é determinada pela norma de reação.

► 18.5 Determinando as Normas de Reação

Sabemos muito pouco sobre as normas de reação de qualquer característica quantitativa em qualquer espécie, em parte porque a determinação de uma norma de reação requer testar muitos membros individuais de genótipo idêntico (ou quase idêntico) em ambientes diferentes. Os animais, entretanto, não são facilmente clonáveis. Uma alternativa a clonar é produzir linhagens de organismos que são idênticos em genótipo por cruzamentos controlados em várias gerações. Como nenhum desses métodos é aplicável em humanos, não

FIG. 18.8 Uma comparação do genótipo *versus* fenótipo em milefólio. (a) *Achillea millefolium*. (b) Norma de reação quanto a elevação para sete plantas *Achillea* diferentes (sete genótipos diferentes). Um broto de cada planta foi cultivado em baixa elevação, um broto em média, e um em alta elevação. [Parte a, Harper Horticultural Slide Library; parte b, Carnegie Institution of Washington.]

temos uma norma de reação para nenhum genótipo para qualquer característica quantitativa humana.

Plantas e animais domésticos

Para determinarmos uma norma de reação, devemos primeiro criar um grupo de indivíduos geneticamente idênticos, uma linhagem homozigota. Esses indivíduos geneticamente idênticos são então deixados desenvolver-se em ambientes diferentes para determinar uma norma de reação. Alternativamente, duas linhagens homozigotas diferentes podem ser cruzadas, e a prole F_1 heterozigota, todos geneticamente idênticos uns aos outros, pode ser testada em ambientes diferentes.

Para muitas plantas, podem ser produzidos clones idênticos pelo simples método de mudas de uma única planta e o cultivo de cada muda em uma planta completa. Um exemplo de normas de reação determinadas por clonagem de plantas é mostrado na Figura 18.8. Plantas individuais de mil-folhas, *Achillea millefolium*, foram coletadas na natureza, e cada planta foi cortada em três pedaços. Um pedaço de cada planta foi plantado em baixa elevação (30 m acima do nível do mar), um pedaço em elevação média (1.400 m) e um em alta elevação (3.050 m). A Figura 18.8b mostra os indivíduos maduros que se desenvolvem dos brotos de sete plantas; cada grupo de três plantas com genótipo idêntico é alinhado verticalmente na figura, para comparação.

Primeiro, notamos um efeito médio do ambiente: em geral, as plantas crescem pouco em elevação média. Esse efeito não é verdadeiro para cada genótipo; entretanto, o broto da planta 4 cresceu melhor na elevação média. Segundo, notamos que nenhum genótipo é incondicionalmente superior em crescimento a todos os outros. A planta 1 mostrou o melhor crescimento nas elevações baixa e alta, mas teve o pior crescimento na elevação média. A planta 6 mostrou o segundo pior crescimento em baixa elevação e o segundo melhor em alta elevação. Novamente, vemos a relação complexa entre genótipo e fenótipo. A Figura 18.9 mostra as normas de reação derivadas dos resultados mostrados na Figura 18.8. Cada genótipo tem uma norma de reação diferente, e as normas se cruzam; logo, não podemos identificar nem o “melhor” genótipo nem o “melhor” ambiente para o crescimento de *Achillea*.

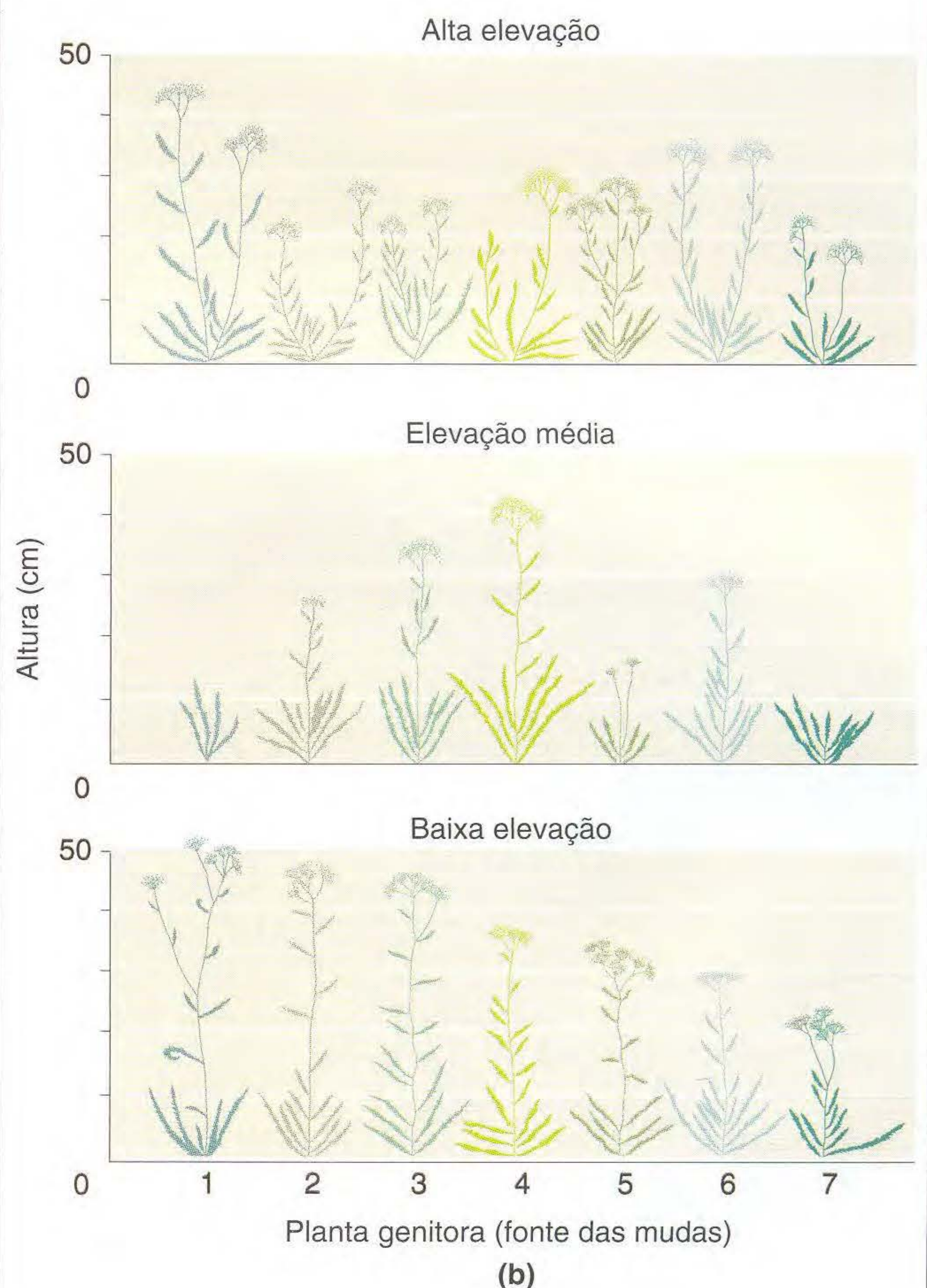
Os genótipos dos organismos de reprodução sexual podem ser replicados pela técnica de cruzar parentes próximos, ou endogamia. Com o uso da autofecundação (quando possível) ou de cruzar irmãos com irmãs repetidamente geração após geração, uma **linhagem segregante** (aquela que contém tanto homozigotos quanto heterozigotos em um locus) pode tornar-se homozigota.

Idealmente, para um estudo de norma de reação, todos os indivíduos devem ser absolutamente idênticos genética-

Os fenótipos podem ser diferentes em ambientes diversos



(a)



mente, mas o processo de endogamia aumenta a homozigose do grupo lentamente, geração após geração, dependendo da proximidade dos parentes que são cruzados. Em milho, por exemplo, uma única planta é escolhida e autofecundada. Então, na geração seguinte, uma de sua prole é escolhida e autofecundada. Na terceira geração, uma única de sua prole é escolhida e autofecundada, e assim por diante. Suponha que a planta original na primeira geração já é homozigota em algum locus. Então, toda a sua prole de autopolinização também será homozigota e idêntica nesse locus. As futuras gerações de autopolinização simplesmente preservarão a homozigose. Se, por outro lado, a planta original for heterozigota, então a autofecundação $A/a \times A/a$ produzirá uma prole que é $\frac{1}{4}$ homozigota A/A e $\frac{1}{4}$ homozigota a/a . Se for escolhida uma única prole para propagar a linhagem, então há uma chance de 50% de que ela agora seja homozigota. Se, por azar, a planta escolhida ainda for heterozigota, existem outros 50% de chance de que a planta selecionada na terceira geração seja homozigota, e assim por diante. Do conjunto de todos os loci heterozigotos, então, após uma geração de autofecundação, apenas 1/2 ainda será heterozigota; após duas gerações, 1/4; após três, 1/8. Na enegésima geração,

$$Het_n = \frac{1}{2^n} Het_0$$

onde Het_n é a proporção dos loci heterozigotos na enegésima geração e Het_0 é a proporção na geração 0. Quando a autofecundação não é possível, a reprodução entre irmão-irmã atinge a mesma finalidade, embora mais lentamente. O Quadro 18.1 é uma comparação da quantidade de heterozigose deixada após n gerações de autofecundação e reprodução entre irmão-irmã.

Estudos de populações naturais

Para fazer um estudo de norma de reação de uma população natural, um grande número de linhagens é colhido da popu-

Quadro 18.1 Heterozigose Restante após Várias Gerações de Endogamia para Dois Sistemas de Reprodução

Geração	Heterozigose restante	
	Autofecundação	Reprodução irmão-irmã
0	1,000	1,000
1	0,500	0,750
2	0,250	0,625
3	0,125	0,500
4	0,0625	0,406
5	0,03125	0,338
10	0,000977	0,114
20	$0,95 \times 10^{-6}$	0,014
n	$Het_n = \frac{1}{2} Het_{n-1}$	$Het_n = \frac{1}{2} Het_{n-1} + \frac{1}{4} Het_{n-2}$

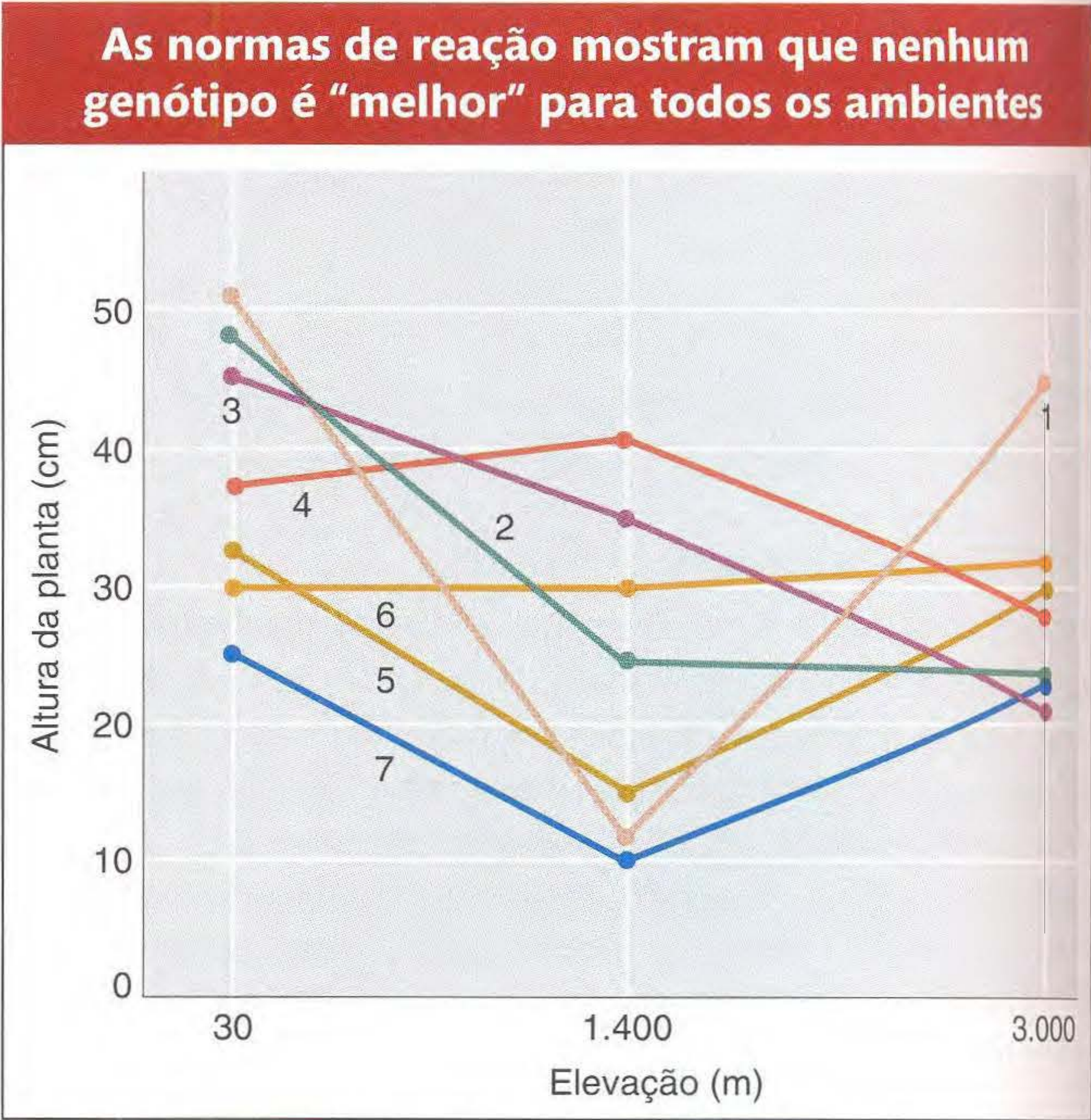


FIG. 18.9 Representação gráfica do conjunto completo de resultados do tipo mostrado na Figura 18.8. Cada linha representa a norma de reação do genótipo de uma planta.

lação e endocruzado por um número suficiente de gerações para garantir que cada linhagem seja de fato homozigota em todos os seus loci. Cada linhagem é então homozigota em cada locus para um alelo selecionado aleatoriamente e presente na população original. As linhagens endocruzadas não podem ser usadas para caracterizar normas de reação na população natural, pois tais genótipos totalmente homozigotos não existem na população original. Cada linhagem endocruzada pode ser cruzada com qualquer outra linhagem endocruzada para produzir heterozigotos que reconstituem a população original, e um número arbitrário de indivíduos de cada cruzamento pode ser produzido. Se a linhagem endocruzada 1 tem a constituição genética $A/A \cdot B/B \cdot c/c \cdot d/d \cdot E/E$ e a linhagem endocruzada 2 é $a/a \cdot B/B \cdot C/C \cdot d/d \cdot e/e$, então um cruzamento entre elas produzirá um grande número de proles, todas as quais são identicamente $A/a \cdot B/B \cdot C/c \cdot d/d \cdot E/e$, e podem ser criadas em ambientes diferentes.

Resultados dos estudos de normas de reação

Muito poucos estudos de normas de reação foram feitos para características quantitativas encontradas em populações naturais, mas muitos foram feitos para espécies domesticadas, tais como milho, que pode ser autopolinizado, ou morangos, que podem ser propagados clonalmente. Os resultados de tais estudos assemelham-se aos dados na Figura 18.9. Nenhum genótipo consistentemente produz um valor fenotípico acima ou abaixo do de outros sob todas as condições ambientais. Ao contrário, existem pequenas diferenças entre os genótipos, e a direção dessas diferenças varia em uma ampla gama de ambientes.

Essas características das normas de reação têm consequências importantes. Uma consequência é que a seleção para genótipos “superiores” em animais domesticados e plantas cultivadas resultará em variedades adaptadas a condições muito específicas, que podem não mostrar suas propriedades superiores em outros ambientes. Até certo ponto, esse problema pode ser superado testando deliberadamente genótipos em uma gama de ambientes (por exemplo, durante vários anos e em vários locais). Entretanto, seria ainda melhor se os horticultores pudessem testar suas seleções em uma variedade de ambientes controlados nos quais os fatores ambientais pudessem ser manipulados separadamente.

As consequências dessas práticas de cruzamentos de plantas podem ser vistas na Figura 18.10, na qual são mostradas as colheitas de duas variedades de milho em função de diferentes ambientes das fazendas. A variedade 1 é uma variedade mais antiga de milho híbrido; a variedade 2 é um híbrido posterior “melhorado”. Seus desempenhos são comparados em uma baixa densidade de plantio, o que era comum quando foi desenvolvida a variedade 1, e em uma alta densidade de plantio, característica das práticas agrícolas quando a variedade 2 foi criada. Na alta densidade de plantio, a variedade 2 é claramente superior à variedade 1 em todos os ambientes (Figura 18.10a). Na baixa densidade de plantio (Figura 18.10b), entretanto, a

situação é bem diferente. Primeiro, note que a nova variedade é menos sensível à variação ambiental que o híbrido mais antigo, como evidenciado por sua norma de reação menos íngreme. Segundo, a nova variedade “melhorada” tem um desempenho mais pobre do que a variedade mais antiga sob as melhores condições agrícolas. Terceiro, a melhora de colheita da nova variedade não ocorre sob baixas densidades de plantio características das antigas práticas agrícolas.

A natureza das normas de reação também têm implicações para as relações sociais humanas e políticas. Mesmo que fique claro que existe variação genética para várias características mentais e emocionais na espécie humana — o que não está claro —, essa variação improvavelmente favorece um genótipo em relação a outro em uma faixa de ambientes. Devemos ter cuidado com normas de reação hipotéticas para características cognitivas humanas que mostram um genótipo sendo incondicionalmente superior a outro. Mesmo deixando de lado todas as questões de julgamento moral e político, simplesmente não existe base para descrever genótipos humanos diferentes como “melhores” ou “piores” em qualquer escala, a menos que o pesquisador seja capaz de fazer uma especificação muito exata do ambiente.

Um genótipo pode ser superior apenas em algumas condições

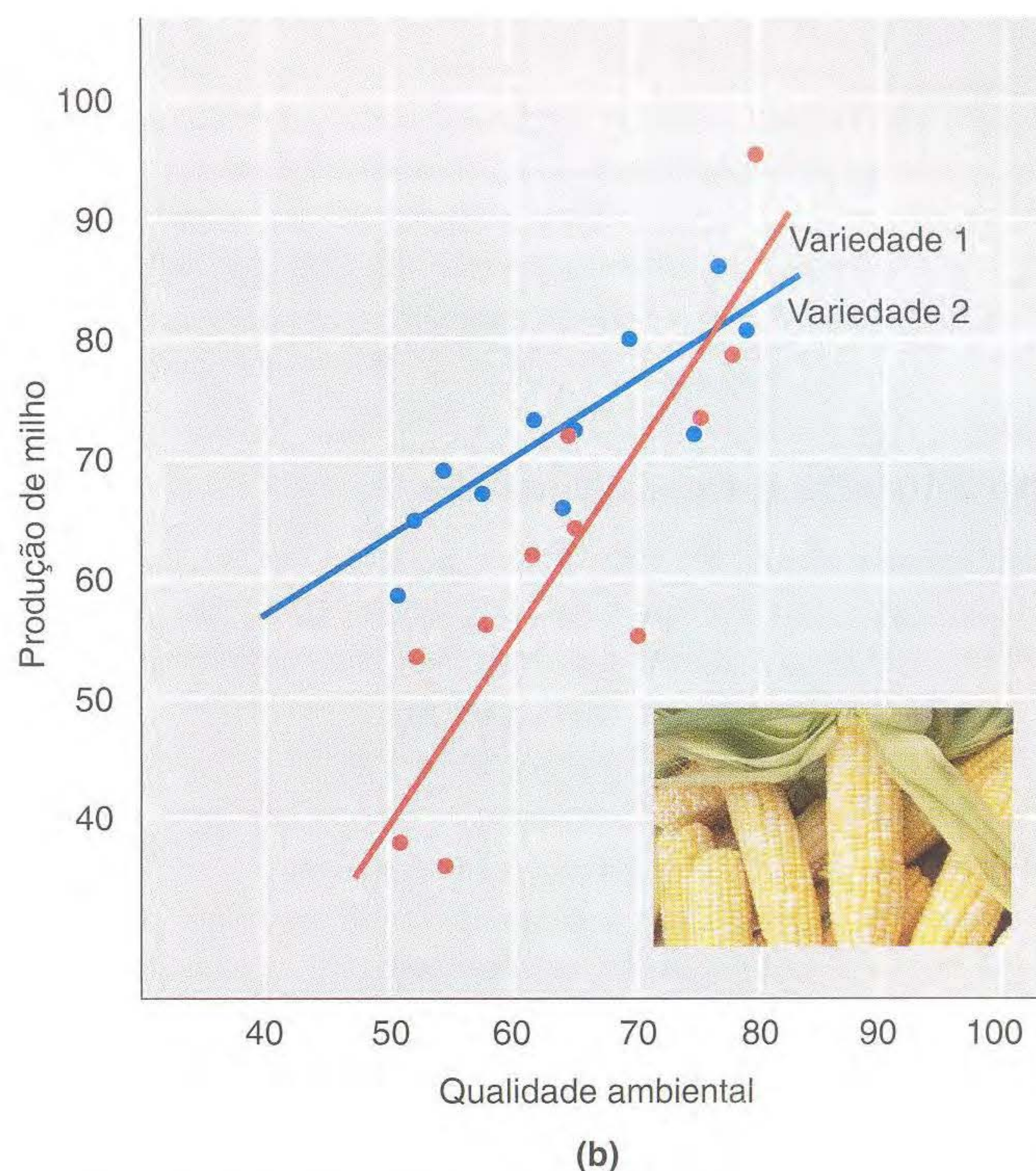
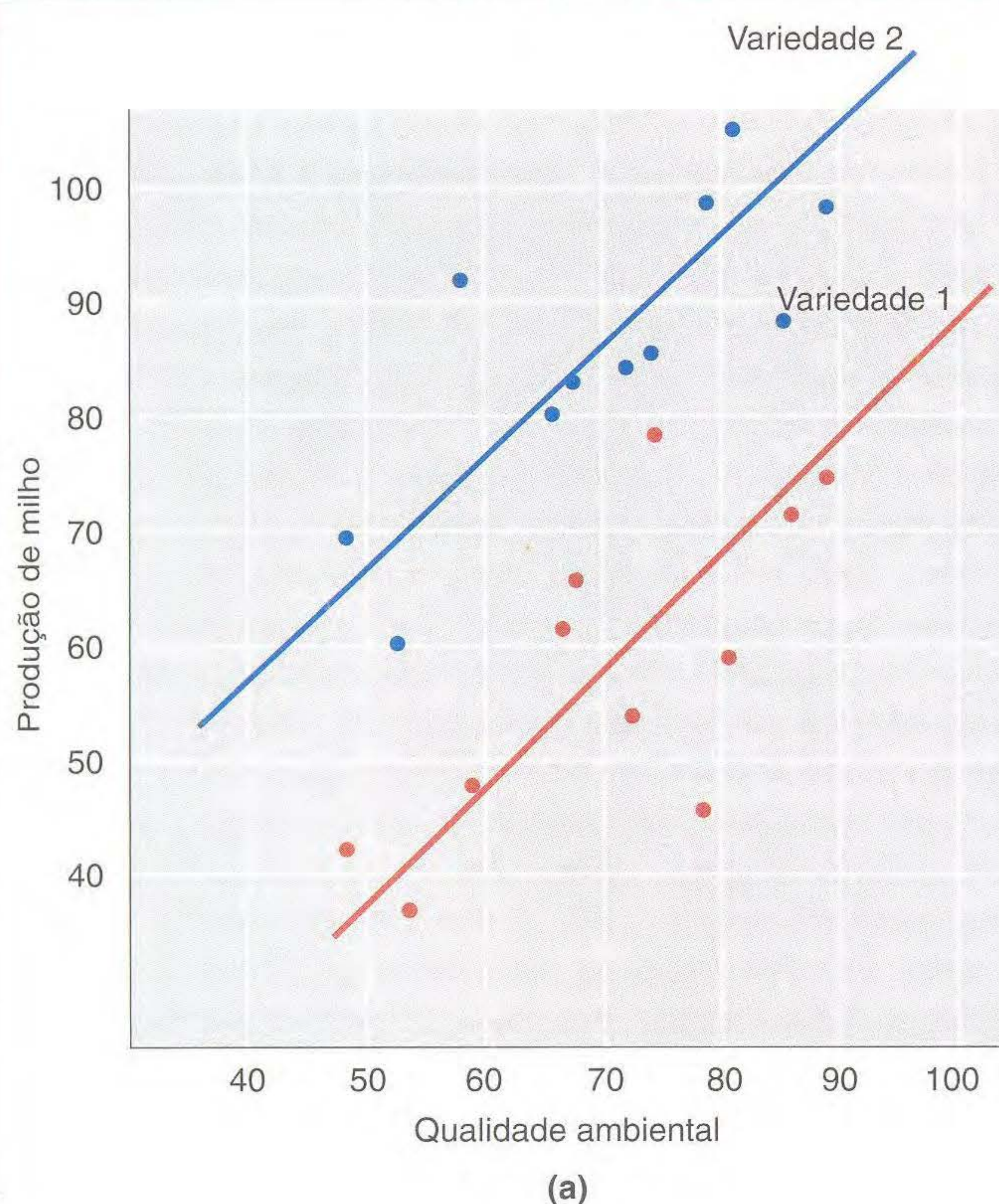


FIG. 18.10 Ambiente e produtividade em grãos. A produtividade em grãos de duas variedades de milho em ambientes diferentes: (a) em uma alta densidade de plantio; (b) em uma baixa densidade de plantio. [Dados por cortesia de W.A. Russell, *Proceedings of the 29th Annual Corn and Sorghum Research Conference*, 1974. Photograph copyright by Bonnie Sue/Photo Researchers.]

Os estudos de normas de reação mostram que, dentro de um único ambiente, existem apenas pequenas diferenças fenotípicas entre a maioria dos genótipos nas populações naturais, e que essas diferenças não são consistentes em uma ampla gama de ambientes. Assim, genótipos "superiores" em animais domésticos e plantas cultivadas podem ser superiores apenas em alguns ambientes. Como nas características físicas, se for demonstrado que os humanos exibem variações genéticas para várias características mentais e emocionais, nenhum genótipo provavelmente supera um outro em uma gama de ambientes.

► 18.6 A Herdabilidade de uma Característica Quantitativa

A pergunta mais básica que podemos fazer sobre uma característica quantitativa é se a variação observada nessa característica é influenciada por genes. É importante notar que essa pergunta não é o mesmo que perguntar se os genes têm algum papel no desenvolvimento da característica. Os processos desenvolvimentais mediados por genes estão na base de toda característica, mas a *variação* em uma característica de indivíduo para indivíduo não é necessariamente o resultado de *variação genética*. Por exemplo, a habilidade em falar uma língua qualquer depende sobretudo das estruturas do sistema nervoso central, bem como das cordas vocais, língua, boca e ouvidos, os quais, por sua vez, dependem de muitos genes no genoma humano. Não existe nenhum ambiente no qual as vacas falem. Mas, embora a língua particular que é falada por humanos varie de nação para nação, essa variação não é genética. Uma característica é dita **herdável** apenas se houver variação genética nessa característica.

A pergunta "Uma característica é herdável?" é uma pergunta sobre o papel que as diferenças nos genes têm nas diferenças fenotípicas entre indivíduos ou grupos.

Familiaridade e herdabilidade

Em princípio, é fácil determinar se alguma variação genética influencia a variação fenotípica em determinada característica. Se os genes têm um papel, então (em média) os parentes biológicos devem ser similares uns aos outros mais do que indivíduos não-aparentados. Essa semelhança seria vista como uma correlação positiva nos valores de uma característica entre genitores e prole, ou entre irmãos (prole dos mesmos genitores). Os genitores que são maiores do que a média, por exemplo, teriam uma prole maior do que a média. Quanto mais sementes uma planta produz, mais sementes suas irmãs produziriam. Tais correlações entre parentes, entretanto, são evidências de variação genética *apenas se os parentes não compartilharem ambientes comuns mais do que os não-parentes fazem*. É absolutamente fundamental distinguir *familiaridade* de *herdabilidade*. Estados de caráter são **familiais** se os membros da mesma família os têm em comum,

por qualquer motivo. Eles só são herdáveis se a similaridade surgir de terem genótipos em comum.

Existem dois métodos gerais para estabelecer a herdabilidade de uma característica como distinta de sua familiaridade. O primeiro depende da *similaridade fenotípica* entre os parentes. Na maioria da história da Genética, esse método tem sido o único disponível, e, assim, quase toda a evidência sobre herdabilidade da maioria das características em organismos experimentais e em humanos foi estabelecida usando esse enfoque. O segundo método, empregando *segregação de genes marcadores*, usa um gene marcador como um tipo de marco para os genes desconhecidos que causam variação em uma característica quantitativa. Um gene marcador é um gene de localização bem conhecida que pode não ter nada a ver com o caráter em estudo. O método depende de mostrar que alelos diferentes de um gene marcador estão associados a fenótipos médios diferentes para o caráter em estudo. Se um gene marcador é visto variando em relação ao caráter, então supostamente ele está ligado a genes que *influenciam* o caráter e sua variação. Assim, a herdabilidade é demonstrada mesmo se os genes que causam a variação no caráter não forem conhecidos. Esse método requer que o organismo em estudo tenha números maiores de loci marcadores geneticamente variáveis e detectáveis dispersos por seu genoma. Tais loci marcadores podem ser observados por variantes na sequência de DNA, estudos eletroforéticos de variação de proteína, ou, nos vertebrados, estudos imunológicos de proteínas de grupos sanguíneos. Dentro de grupos, por exemplo, galinhas com grupos sanguíneos diferentes apresentam alguma diferença no peso dos ovos, mas, tanto quanto se sabe, os antígenos de grupos sanguíneos e anticorpos não causam, em si, a diferença no tamanho do ovo. Supostamente, os genes que influenciam o peso do ovo estão ligados a loci que determinam o grupo sanguíneo.

Desde a introdução dos métodos moleculares para o estudo das sequências de DNA, uma grande quantidade de variação genética foi descoberta em uma grande variedade de organismos. Essa variação consiste ou em substituições em posição de um único nucleotídeo ou de números variáveis de inserções ou repetições de seções curtas de DNA. Essas variações são geralmente detectadas pelo ganho ou perda de sítios de reconhecimento de enzimas de restrição, ou pelo tamanho da variação nas sequências de DNA entre dois sítios de restrição fixos, os quais são formas de polimorfismo de comprimento de fragmento de restrição (RFLP; veja o Capítulo 17). Em tomates, por exemplo, as linhagens que levam variantes diferentes de RFLP diferem nas características dos frutos. É suposto que as sequências de DNA nesses RFLP não influenciam em si as características do fruto. Eles são marcos situados perto de genes que, portanto, apresentam altos níveis de co-segregação para essas características.

Como muito o que se sabe ou é dito sobre a herdabilidade ainda depende de similaridade fenotípica entre parentes, entretanto, sobretudo em genética humana, começaremos nosso exame do problema de herdabilidade estudando a similaridade fenotípica.

Similaridade fenotípica entre parentes

Em organismos experimentais, não há problema em separar as similaridades ambientais das genéticas. A prole de uma vaca produtora de uma grande quantidade de leite e a prole de uma vaca produtora de pouco leite podem ser criadas juntas no mesmo ambiente para ver se, a despeito da similaridade ambiental, cada uma se assemelha à sua genitora. Em populações naturais, entretanto, e especialmente em humanos, esse tipo de estudo é difícil de fazer. Devido à natureza das sociedades humanas, os membros da mesma família não só têm genes em comum, mas também ambientes similares. Assim, a observação de simples similaridade familiar de fenótipo é geneticamente não-interpretável. Em geral, as pessoas que falam húngaro têm genitores que falam húngaro, e pessoas que falam japonês têm genitores falando japonês. Embora a intensa experiência de imigração para a América do Norte tenha demonstrado isso, essas diferenças lingüísticas, conquanto familiares, não são genéticas. As mais altas correlações entre pais e prole para qualquer característica social nos Estados Unidos são aquelas para partidos políticos e filiações religiosas, mas essas características não são herdáveis. A distinção entre familiaridade e hereditariedade, entretanto, nem sempre é tão óbvia. A U.S. Public Health Commission, quando estudou a doença pelagra, por deficiência de vitamina, no Sudeste dos EUA em 1910, chegou à conclusão de que ela era genética porque ocorria nas famílias. Entretanto, a pelagra hoje é bem compreendida como tendo sido prevalente nas populações do Sudeste dos EUA devido a uma dieta pobre.

Para determinarmos se uma característica é herdável, devemos usar estudos de algumas pessoas adotadas para evitar a similaridade ambiental usual entre parentes biológicos. As pessoas experimentalmente ideais são gêmeos monozigóticos (idênticos) criados separadamente, pois eles são geneticamente idênticos, mas vivem em ambientes diferentes. Tais estudos de adoção devem ser feitos de modo que não exista correlação entre os ambientes sociais das famílias que adotam e as famílias biológicas. De outro modo, as similaridades entre os ambientes de gêmeos não teriam sido eliminadas pela adoção. Esses requisitos são muito difíceis de atingir; portanto, na prática, sabemos muito pouco sobre se as características quantitativas humanas que são familiares também são herdáveis.

A cor da pele é claramente herdável, como a altura do adulto, mas, mesmo para características como essas, temos que ter cuidado. Sabemos que a cor da pele é afetada por genes, tanto de estudos de adoções inter-raciais quanto de observações de que a prole de escravos negros africanos era composta de negros, mesmo quando estes eram nascidos e criados nos EUA. Mas as diferenças de altura entre japoneses e europeus são afetadas por genes? Os filhos de imigrantes japoneses que nasceram e foram criados na América do Norte são mais altos do que seus genitores, porém mais baixos do que a média de norte-americanos, e, assim, podemos concluir que há alguma influência de diferença genética. Entretanto, a segunda geração de nipo-americanos é ainda constituída de indivíduos mais altos que seus genitores nascidos ame-

ricanos. Parece que alguma influência ambiental-cultural, possivelmente nutricional ou, talvez, um efeito de herança materna (veja o Capítulo 12), ainda é sentida na primeira geração de nascidos nos EUA. Ainda não podemos dizer nada de definitivo sobre as diferenças genéticas que podem contribuir para as diferenças de altura entre norte-americanos de, digamos, ancestrais japoneses e suecos.

Características de personalidade, temperamento, desempenho cognitivo (incluindo valores de QI) e toda uma variedade de comportamentos, como alcoolismo e distúrbios mentais, tais como esquizofrenia, têm sido objeto de estudos de herdabilidade em populações humanas. Muitas dessas características apresentam familiaridade (isto é, similaridade familiar). Há uma correlação positiva, por exemplo, entre valores de QI dos genitores e os valores de seus filhos (a correlação é de cerca de 0,5 nas famílias brancas americanas), mas essa correlação não distingue a familiaridade da herdabilidade. Fazer essa distinção requer que a correlação ambiental entre genitores e filhos seja quebrada, e, assim, os estudos de crianças adotadas são comuns. Como é difícil tornar aleatórios os ambientes, mesmo em casos de adoção, a evidência da herdabilidade para personalidade humana e características comportamentais permanece equívoca a despeito do grande número de estudos existentes. Preconceitos sobre as causas das diferenças humanas são generalizados e profundos, e, como resultado, as evidências associadas aos estudos da herdabilidade do QI, por exemplo, têm sido mais frouxas que os estudos de produção de leite em vacas.

A Figura 18.11 resume o método usual de teste de herdabilidade em organismos experimentais. Os indivíduos de ambos os extremos da distribuição fenotípica são cruzados com outros indivíduos de seu próprio grupo extremo, e a prole é criada em um ambiente comum controlado. Se houver uma diferença média entre os dois grupos de prole, a diferença fenotípica é herdável. A maioria das características morfológicas em *Drosophila*, por exemplo, é herdável, mas nem todas elas. Se as moscas com asas direitas um pouco mais longas do que as esquerdas são cruzadas entre si, sua prole não tem maior tendência para ser “asa direita” do que a prole de moscas “asa esquerda”. Como veremos, esse método também pode ser usado para obter informação quantitativa sobre a herdabilidade.

Em organismos experimentais, as similaridades de ambientes em geral podem ser prontamente distintas da similaridade genética (ou herdabilidade). Em humanos, entretanto, em geral é muito difícil determinar se uma certa característica é herdável.

► 18.7 Quantificando a Herdabilidade

Se uma característica é mostrada como sendo herdável em uma população, então é possível quantificar o grau de herdabilidade. As Figuras 18.4 e 18.5 mostram que a variação entre fenótipos surge de duas fontes: primeira, existem

Os cruzamentos são feitos para testar a herdabilidade

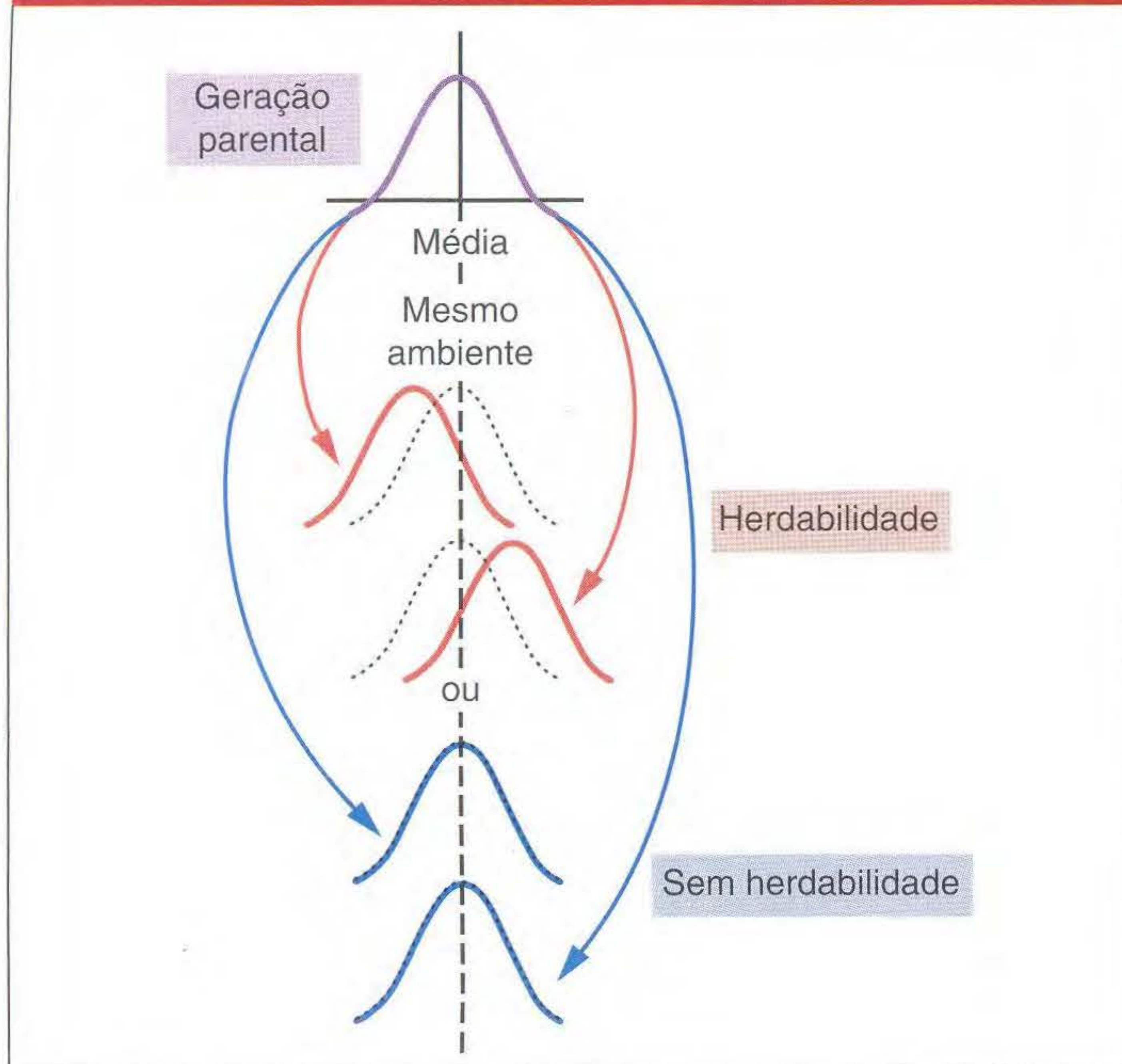


FIG. 18.11 Método padrão para testar herdabilidade em organismos experimentais. Os cruzamentos são feitos dentre duas populações de indivíduos selecionados pelos extremos de distribuição fenotípica na geração parental. Se as distribuições fenotípicas dos dois grupos de prole são significativamente diferentes uma da outra (curvas vermelhas), então a diferença de característica é herdável. Se ambas as distribuições de prole se assemelham à distribuição na geração parental (curvas azuis), então a diferença fenotípica não é herdável.

diferenças médias de genótipos; segunda, cada genótipo exibe variação fenotípica devido à variação ambiental. A variação fenotípica total da população (s_p^2) pode assim ser fracionada em duas partes: a variação entre as médias genotípicas (s_g^2) e a variação restante (s_e^2). A primeira é chamada de **variação genética**, e a última é chamada de **variação ambiental**; entretanto, como veremos, esses nomes são muito enganadores. Além disso, a quebra da variação fenotípica em variações ambientais e genéticas deixa a possibilidade de alguma covariância entre genótipo e ambiente. Por exemplo, suponha que seja verdade (não sabemos isso) que existem genes que influenciam a habilidade musical em humanos. Os genitores com tais genes podem ser músicos, que criariam um ambiente mais musical para seus filhos, que então teriam tanto genes quanto ambiente promovendo o desempenho musical. O resultado seria uma variação maior entre pessoas quanto à habilidade musical do que seria o caso se não houvesse efeito do ambiente parental nas crianças. Se o fenótipo é a soma de um efeito genético e um efeito ambiental, $p = g + e$, então a variação do fenótipo é, de acordo com a fórmula $s_T^2 = s_x^2 + s_y^2 + 2 \text{ cov } xy$, a soma da variação genética, a variação ambiental e o dobro da covariância entre os efeitos genotípicos e os efeitos ambientais:

$$s_p^2 = s_g^2 + s_e^2 + 2 \text{ cov } ge$$

Se os genótipos não são distribuídos aleatoriamente pelos ambientes mas esse fato não é levado em conta, haverá alguma covariância entre os valores genotípicos e ambientais, e essa covariância será escondida nas variações genéticas e ambientais.

A medida quantitativa da herdabilidade de uma característica é a parte da variação fenotípica que é devida à variação genética:

$$H^2 = \frac{s_g^2}{s_p^2} = \frac{s_g^2}{s_g^2 + s_e^2}$$

H^2 , assim definida, é chamada de **herdabilidade ampla** da característica.

Deve ser destacado que essa medida da “influência genética” nos diz que parte da *variação* da população no fenótipo pode ser atribuída à *variação* no genótipo. Ela não nos diz que partes de um fenótipo *individual* podem ser atribuídas a seu genótipo e a seu ambiente. Esta última distinção não é razoável. Um fenótipo de um indivíduo é consequência da interação de seus genes e a seqüência de ambientes que ele experimenta à medida que se desenvolve. Seria tolice dizer que 1,5 m de sua altura foi produzido por seus genes e 25 cm foram produzidos pelo ambiente. Todas as medidas da “importância” dos genes são formuladas em termos da proporção da variação fenotípica atribuída à sua variação. Esse enfoque é uma aplicação especial da técnica mais geral de **análise de variância**, usada para atribuir a altura relativa a genes contribuintes. A técnica foi, de fato, originalmente inventada para lidar com experimentos nos quais diferentes fatores ambientais e genéticos estavam influenciando o crescimento das plantas. (Para um tratamento sofisticado porém acessível da análise de variância escrita para biólogos, veja R. Sokal and J. Rohlf, *Biometry*, 3rd ed., W.H. Freeman and Company, 1995.)

Métodos para estimar H^2

A herdabilidade em uma população pode ser estimada de vários modos. Mais diretamente, podemos obter uma estimativa da variação ambiental em uma população, s_e^2 , fazendo várias linhagens homozigotas, cruzando-as aos pares para fazer heterozigotos típicos da população e medindo a variação fenotípica *dentro* de cada genótipo heterozigoto. Como todos os indivíduos dentro de cada grupo têm o mesmo genótipo, e, portanto, não há variação genética dentro dos grupos, essas variações irão (em média) fornecer uma estimativa de s_e^2 . Esse valor pode então ser subtraído do valor de s_p^2 na população original para dar s_g^2 . Com o uso desse método, qualquer covariância entre genótipo e ambiente na população original será escondida na estimativa da variação genética e irá inflá-la. Assim, por exemplo, se os indivíduos com genótipos que os fariam mais altos em média em ambientes aleatórios também recebessem melhor nutrição que indivíduos com genótipos que os fariam mais baixos, então a diferença observada nas alturas entre os dois grupos genotípicos seria exagerada.

Outras estimativas de variação genética podem ser obtidas considerando as similaridades genéticas entre parentes.

Usando simples princípios mendelianos, podemos ver que metade dos genes de irmãos serão (em média) idênticos. Por simplificação, podemos marcar os alelos em um locus levados apenas pelos genitores, digamos, como A_1/A_2 e A_3/A_4 . Os irmãos mais velhos têm a probabilidade de $1/2$ de receber A_1 de seu pai, como o irmão mais novo, e, assim, dois irmãos têm uma chance de $1/2 \times 1/2 = 1/4$ de ambos terem A_1 . Por outro lado, eles podem ambos receber A_2 de seu pai; e, assim, novamente, têm a probabilidade de $1/4$ de levar esse alelo. Assim, a chance é $1/4 + 1/4 = 1/2$ de que ambos os irmãos herdem o mesmo alelo (seja A_1 ou A_2) de seu pai. Na outra metade das vezes, um irmão herdará um A_1 e o outro herdará um A_2 . Assim, no que se refere a genes herdados paternamente, os irmãos têm uma chance de 50% de levar o mesmo alelo. Mas o mesmo raciocínio aplica-se a seu alelo herdado maternamente. Na média entre seus alelos herdados paterna e maternamente $[(1/2 + 1/2)/2 = 1/2]$, metade dos genes de irmãos serão idênticos entre eles. Sua **correlação genética**, que é igual à chance de que levem o mesmo alelo, será de $1/2$, ou 0,5.

Se aplicarmos esse raciocínio a meios-irmãos, com um pai comum mas com mães diferentes, teremos um resultado diferente. Novamente, os dois irmãos têm 50% de chance de herdar um gene idêntico de seu pai, mas, dessa vez, não há modo de herdar o mesmo gene de suas mães, pois têm duas mães diferentes. Na média dos genes herdados materna e paternamente, temos uma probabilidade de $(1/2 + 0)/2 = 1/4$ de que esses meios-irmãos levem o mesmo gene.

Podemos ser tentados a usar essa correlação teórica entre parentes para estimar H^2 . Se a correlação fenotípica observada entre irmãos fosse, por exemplo, 0,4, e esperássemos, apenas do ponto de vista genético, uma correlação de 0,5, então nossa estimativa de herdabilidade seria $0,4/0,5 = 0,8$. Mas tal estimativa falha em levar em conta o fato de que os ambientes dos irmãos também podem estar correlacionados. A menos que tenhamos cuidado em criar os irmãos em ambientes independentes, nossa estimativa de H^2 seria muito grande e podia exceder 1 se a correlação fenotípica fosse maior que 0,5.

Para contornar esse problema, usamos as *diferenças* entre correlações de parentes diferentes. Por exemplo, a diferença em correlação genética entre irmãos e meios-irmãos é $1/2 - 1/4 = 1/4$. Vamos contrastar isso com suas **correlações fenotípicas**. Se a similaridade ambiental é a mesma para irmãos e meios-irmãos, uma condição muito importante para estimar a herdabilidade, então as similaridades ambientais serão canceladas se tomarmos a diferença em correlação entre os dois tipos de irmãos. Essa diferença na correlação fenotípica será então proporcional ao quanto da variância é genética. Assim:

$$\left(\begin{array}{c} \text{correlação genética} \\ \text{de irmãos} \end{array} \right) - \left(\begin{array}{c} \text{correlação genética} \\ \text{de meios-irmãos} \end{array} \right) = \frac{1}{4}$$

mas

$$\left(\begin{array}{c} \text{correlação fenotípica} \\ \text{de irmãos} \end{array} \right) - \left(\begin{array}{c} \text{correlação fenotípica} \\ \text{de meios-irmãos} \end{array} \right) = H^2 \times \frac{1}{4}$$

e, assim, uma estimativa de H^2 é

$$H^2 = 4 \left[\left(\begin{array}{c} \text{correlação} \\ \text{de irmãos} \end{array} \right) - \left(\begin{array}{c} \text{correlação de} \\ \text{meios-irmãos} \end{array} \right) \right]$$

onde a correlação aqui é a correlação *fenotípica*.

Essa estimativa, bem como outras baseadas em correlações entre parentes, depende *criticamente* da suposição de que as correlações ambientais entre indivíduos são as mesmas para todos os graus de correlação, o que não é provável que seja o caso. Os irmãos, por exemplo, são geralmente criados pelo mesmo par de genitores, enquanto os meios-irmãos provavelmente são criados em circunstâncias apenas com um genitor em comum. Se os parentes próximos têm mais ambientes similares, como entre os humanos, essas estimativas de herdabilidade serão distorcidas. É razoável supor que a maioria das correlações ambientais entre parentes é positiva, em cujo caso as herdabilidades serão superestimadas. Mas as correlações ambientais negativas também podem existir. Por exemplo, se os membros de uma prole têm que competir por comida, que é pouca, pode haver correlações negativas nas taxas de crescimento entre os irmãos.

A diferença na correlação fenotípica entre gêmeos mono- zigóticos e dizigóticos é comumente usada em genética humana para estimar H^2 quanto a características cognitivas ou de personalidade. Aqui, o problema do grau de similaridade ambiental é muito grave. Os gêmeos mono- zigóticos (idênticos) são geralmente tratados mais similarmente um com o outro do que os gêmeos dizigóticos (fraternos). Os genitores geralmente dão aos gêmeos idênticos nomes que são similares, vestem-nos de modos similares, tratam-nos do mesmo modo e, em geral, acentuam suas similaridades. Como resultado, a herdabilidade é superestimada.

O significado de H^2

A atenção aos problemas de estimar a herdabilidade ampla diverge de questões mais profundas sobre o significado da proporção mesmo quando ela pode ser estimada. A despeito do uso generalizado como uma medida do quanto genes “importantes” influenciam uma característica, H^2 tem um significado especial e limitado.

Duas conclusões alternativas podem ser tiradas dos resultados de um estudo de herdabilidade apropriadamente elaborado. Primeiro, se há uma herdabilidade não-zero, podemos concluir que, na população medida e nos ambientes nos quais os organismos se desenvolveram, as diferenças genéticas influenciaram a variação fenotípica entre os indivíduos, e, assim, as diferenças genéticas são importantes para a característica. Esse achado não é trivial, e é a primeira etapa em uma investigação mais detalhada do papel dos genes.

É importante notar que o reverso não é verdadeiro. Se forem encontradas herdabilidade zero para a característica, esse achado não é uma demonstração de que os genes são irrelevantes para a característica; ao contrário, ele demonstra apenas que, em determinada população e ambiente estudado, ou não há variação genética em loci relevantes, ou genótipos diferentes têm o mesmo fenótipo. Em outras populações ou em outros ambientes, a característica pode ser herdável.

A herdabilidade de uma diferença de característica é diferente em cada população e em cada conjunto de ambientes; ela não pode ser extrapolada de uma população e conjunto de ambientes para outro.

Além disso, devemos distinguir entre genes que contribuem para uma característica e *diferenças genéticas* que contribuem para *diferenças* em uma característica. O experimento natural de imigração para os EUA mostrou que a habilidade em pronunciar os sons do inglês norte-americano, em vez do francês, sueco ou russo, não é uma consequência de diferenças genéticas entre nossos ancestrais imigrantes. Mas, sem os genes apropriados, não poderíamos falar língua nenhuma.

Segundo, o valor de H^2 dá uma previsão limitada de como uma característica pode ser modificada trocando o ambiente. Se toda a variação ambiental relevante é eliminada e o novo ambiente constante é o mesmo que o ambiente médio na população original, então H^2 estima o quanto da variação fenotípica ainda estará presente. Assim, se a herdabilidade do desempenho em um teste de QI fosse encontrado como sendo, digamos, 0,4, então poderíamos prever que, se todas as crianças tivessem o mesmo ambiente desenvolvimental e social que a “criança média”, cerca de 60% da variação no teste de QI iriam desaparecer e 40% permaneceriam.

O requisito de que o novo ambiente constante seja uma média da distribuição ambiental antiga é absolutamente essencial para essa previsão. Se o ambiente é mudado para uma extremidade ou outra da distribuição ambiental presente na população usada para determinar H^2 , ou se um novo ambiente é introduzido, nada pode ser previsto. No exemplo do teste de desempenho de QI, a herdabilidade não nos dá nenhuma informação sobre o quanto o desempenho seria variável se os ambientes desenvolvimentais e sociais de todas as crianças fossem enriquecidos. Para compreendermos por que isso é assim, devemos retornar ao conceito de norma de reação.

A separação da variância fenotípica em componentes genético e ambiental, s_g^2 e s_e^2 , de fato não separa as causas ambientais e genéticas da variação. Considere os resultados apresentados na Figura 18.10b. Quando o ambiente é pobre (uma qualidade ambiental de 50), a variedade 2 de milho tem uma produtividade maior que a variedade 1, e, assim, uma população feita de uma mistura das duas variedades teria muita variância genética para produzir nesse ambiente. Mas, em um ambiente mais rico (valor 75), não há diferença de produtividade entre as variedades 1 e 2, e, assim, uma população mista não teria nenhuma variância genética para produzir nesse ambiente. Assim, a variância genética foi mudada trocando o ambiente. Por outro lado, a variedade 2 é menos sensível ao ambiente que a variedade 1, como mostrado pelas inclinações das duas linhas. Assim, uma população feita principalmente da variedade 2 teria uma menor variância ambiental do que uma feita principalmente da variedade 1. Assim, a variância ambiental na população é mudada trocando a proporção de genótipos.

Como consequência do argumento apresentado, podemos prever, a partir do conhecimento da herdabilidade da diferença de uma característica, como a distribuição da variação na característica irá mudar se as frequências genóticas ou fatores ambientais mudarem acentuadamente. Assim, por exemplo, com relação ao desempenho no teste de QI, saber que a herdabilidade é 0,4 em um ambiente não nos permite prever como o teste de QI irá variar entre crianças em um ambiente diferente.

Uma alta herdabilidade não significa que uma característica não é afetada pelo ambiente. Como o genótipo e o ambiente interagem para produzir o fenótipo, nenhuma partição da variação em seus componentes genéticos e ambientais pode de fato separar as causas da variação.

Toda essa alta herdabilidade significa que, para o desenvolvimento particular de uma população na distribuição particular dos ambientes nos quais a herdabilidade foi medida, as diferenças médias entre genótipos são grandes comparadas com a variação ambiental dentro dos genótipos. Se o ambiente é mudado, então pode haver grandes diferenças em fenótipo.

Talvez o exemplo mais conhecido do uso errado dos argumentos de herdabilidade para falar sobre mutabilidade de uma característica seja o desempenho do QI humano é o sucesso social. Muitos estudos foram feitos sobre a herdabilidade do desempenho do QI na crença de que, se a herdabilidade é alta, então vários programas de educação destinados a aumentar o desempenho intelectual são uma perda de tempo. O argumento é que, se uma característica é altamente herdável, então ela não pode ser muito mudada por alterações ambientais. Mas, a despeito do valor de H^2 para o desempenho no teste de QI, o erro real do argumento está em igualar a alta herdabilidade com imutabilidade. De fato, a herdabilidade do QI é irrelevante para a questão do quanto ela é mutável.

Para vermos o porquê disso, consideremos os resultados usuais dos estudos de QI em crianças que foram separadas de seus genitores biológicos na lactância e criados por pais adotivos. Embora esses resultados variem quantitativamente de estudo para estudo, eles têm três características em comum. Primeiro, como os pais adotivos geralmente vêm de uma população mais bem educada do que os pais biológicos que dão seus filhos para adoção, eles geralmente têm valores mais altos de QI do que os pais biológicos. Segundo, os filhos adotados têm valores de QI maiores do que os de seus genitores biológicos. Terceiro, os filhos adotados mostram uma correlação maior de valores de QI com seus pais biológicos do que com seus pais adotivos. O quadro seguinte é um conjunto de dados hipotéticos que mostram todas essas características, de forma idealizada, para ilustrar esses conceitos. Os valores dados para genitores pretendem ser a média entre a mãe e o pai.

	Filhos	Pais biológicos	Pais adotivos
	110	90	118
	112	92	114
	114	94	110
	116	96	120
	118	98	112
	120	100	116
Média	115	95	115

Primeiro, podemos ver que os valores das crianças têm uma correlação maior com os de seus pais biológicos, mas uma correlação baixa com os de seus pais adotivos. De fato, em nosso exemplo hipotético, a correlação de crianças com os pais biológicos é $r = 1,00$, mas com os pais adotivos é $r = 0$. (Lembre que uma correlação entre dois conjuntos de números não significa que os dois conjuntos sejam idênticos, mas que, para cada unidade de aumento em um conjunto, há um constante aumento proporcional no outro conjunto; veja Apêndice Estatístico na análise estatística no final deste capítulo.) Essa perfeita correlação com pais biológicos e correlação zero com pais adotivos significa que $H^2 = 1$, considerando os argumentos desenvolvidos. Toda a variação de QI entre as crianças é explicada pela variação de QI entre pais biológicos, que não tiveram chance de influenciar os ambientes de seus filhos.

Segundo, entretanto, notamos que o valor do QI de cada criança é 20 pontos mais alto do que os de seus pais biológicos, e que o QI médio das crianças é igual ao QI médio dos genitores adotivos. Assim, a adoção aumentou o QI médio das crianças 20 pontos acima da média de QI de seus pais biológicos, e, desse modo, como um grupo, as crianças assemelham-se a seus pais adotivos. Assim, temos uma herdabilidade perfeita, embora alta plasticidade em resposta à modificação ambiental.

Um pesquisador seriamente interessado em saber como os genes podem restringir ou influenciar o curso do desenvolvimento de qualquer característica em qualquer organismo deve estudar diretamente as normas de reação dos vários genótipos na população na gama de ambientes projetados. Nenhuma informação menos detalhada o faria. Medidas resumidas, tais como H^2 , não são valiosas em si.

A herdabilidade não é o oposto de plasticidade fenotípica. Uma característica pode ser perfeitamente herdável em uma população e ainda estar sujeita a grandes mudanças resultantes de variação ambiental.

Herdabilidade restrita

O conhecimento da herdabilidade ampla (H^2) de uma característica em uma população não é muito útil em si, mas uma subdivisão mais fina da variação fenotípica pode das informações importantes para os criadores de animais e plantas. A

variação genética pode em si ser subdividida em dois componentes para da informações sobre a ação gênica e a possibilidade de moldar a composição genética de uma população.

Nossa consideração prévia da ação gênica sugere que os fenótipos dos homozigotos e heterozigotos têm que ter uma relação simples. Se um dos alelos codifica um produto gênico menos ativo ou um sem nenhuma atividade, e se uma unidade do produto gênico for suficiente para permitir a total atividade fisiológica do organismo, então esperaríamos a completa dominância de um alelo sobre o outro, como Mendel observou para cor das flores em ervilhas. Se, por outro lado, a atividade fisiológica fosse proporcional à quantidade do produto gênico ativo, esperaríamos que o fenótipo heterozigoto fosse exatamente intermediário aos homozigotos (não apresentando dominância).

Para muitas características quantitativas, entretanto, nenhum desses casos simples é a regra. Em geral, os heterozigotos não são exatamente intermediários dos dois homozigotos, mas são próximos de um ou do outro (apresentam dominância parcial), muito embora exista uma mistura igual de produtos primários dos dois alelos no heterozigoto. Suponha que dois alelos, a e A , segreguem em um locus influenciando a altura. Nos ambientes encontrados pela população, os fenótipos médios (alturas) e frequências dos três genótipos podem ser:

	a/a	A/a	A/A
Fenótipo	10	18	20
Frequência	0,36	0,48	0,16

Há uma variação genética na população; as médias fenotípicas das três classes genotípicas são diferentes. Parte da variação surge porque há um efeito médio no fenótipo de substituir um alelo A por um alelo a ; isto é, a altura média de todos os indivíduos com alelos A é maior do que a de todos os indivíduos com alelos a . Definindo o efeito médio de um alelo como o fenótipo médio de todos os indivíduos que o levam, necessariamente fazemos com que o efeito médio do alelo dependa das frequências dos genótipos.

O efeito médio é calculado simplesmente contando os alelos a e A e multiplicando-os pelas alturas dos indivíduos nos quais eles aparecem. Assim, 0,36 de todos os indivíduos são homozigotos a/a , cada indivíduo a/a tendo dois alelos a , e a altura média dos indivíduos a/a é de 10 cm. Os heterozigotos constituem 0,48 da população, cada um com um alelo a , e a medida fenotípica média dos indivíduos A/a é 18 cm. O “número” total de alelos a é $2(0,36) + 1(0,48)$. Assim, o efeito médio de todos os alelos a é

$$\begin{aligned}\bar{a} &= \text{efeito médio de } a = \frac{2(0,36)(10) + 1(0,48)(18)}{2(0,36) + 1(0,48)} \\ &= 13,20 \text{ cm}\end{aligned}$$

e, por argumento similar,

$$\begin{aligned}\bar{A} &= \text{efeito médio de } A = \frac{2(0,16)(20) + 1(0,48)(18)}{2(0,16) + 1(0,48)} \\ &= 18,80 \text{ cm}\end{aligned}$$

A diferença no efeito médio entre A e a é $18,80 - 13,20 = 5,60$ cm. A diferença no efeito médio é chamada de **efeito aditivo**. Ela responde por parte da variação no fenótipo, mas não por toda ela. O heterozigoto não é exatamente intermediário entre os homozigotos; existe alguma dominância.

Gostaríamos de separar o chamado efeito aditivo causado pela substituição dos alelos a pelos alelos A da variação causada pela dominância. O motivo é que o efeito do cruzamento seletivo depende da variação aditiva, e não da variação causada pela dominância. Assim, para fins de cruzamentos entre plantas ou animais ou para fazer previsões sobre a evolução por seleção natural, devemos determinar a variação aditiva. Um exemplo extremo irá ilustrar o princípio. Suponha que a altura da planta é influenciada por variação em um gene e que os significados fenotípicos e frequências dos três genótipos são:

	A/A	A/a	a/a
Fenótipo	10	12	10
Frequência	0,25	0,50	0,25

É aparente (e um cálculo como o precedente irá confirmar) que não existe diferença média entre os alelos a e A , porque cada um tem um efeito de 11 unidades. Assim, não existe variação *aditiva*, embora exista uma variação genética óbvia, pois há uma variação no fenótipo entre os genótipos. As plantas mais altas são heterozigotas. Se um criador tentar aumentar a altura de sua população por cruzamento seletivo, o cruzamento desses heterozigotos juntos simplesmente reconstituirá a população original. A seleção será totalmente inefetiva. Esse exemplo ilustra a lei geral de que o efeito da seleção depende da variação genética *aditiva*, e não da variação genética em geral.

A variação genética total em uma população pode ser subdividida em dois componentes: **variação genética aditiva** (s_a^2), a variação que surge porque há uma diferença média entre os portadores dos alelos a e os portadores dos alelos A , e a **variação de dominância** (s_d^2), a variação que resulta do fato de que os heterozigotos não são exatamente intermediários aos monozigotos. Assim:

$$s_g^2 = s_a^2 + s_d^2$$

A variação fenotípica total agora pode ser escrita como

$$s_p^2 = s_g^2 + s_e^2 = s_a^2 + s_d^2 + s_e^2$$

Definimos um novo tipo de herdabilidade, a **herdabilidade no sentido restrito** (h^2), como

$$h^2 = \frac{s_a^2}{s_p^2} = \frac{s_a^2}{s_a^2 + s_d^2 + s_e^2}$$

É essa herdabilidade — não confundir com H^2 — que é útil em determinar se o programa de cruzamento seletivo terá sucesso em mudar a população. Quanto maior for h^2 , maior a fração da diferença entre os genitores selecionados e a população como um todo que será preservada na prole dos genitores selecionados.

O efeito da seleção depende da quantidade da variação genética *aditiva* e não da variação genética em geral. Portanto, é a herdabilidade restrita, h^2 , não a herdabilidade ampla, H^2 , que prevê a resposta à seleção.

Estimando os componentes da variação genética

Os componentes diferentes da variação genética podem ser estimados pela covariância entre parentes, o grau pelo qual os fenótipos de pares de parentes são correlacionados uns com os outros, mas o desvio dessas estimativas está além do escopo deste livro. Existe, entretanto, um outro modo de estimar a herdabilidade restrita, h^2 , que revela seu significado real. Se, em duas gerações de uma população, plotamos o fenótipo, digamos altura, da prole contra o fenótipo médio de seus genitores (o **valor médio parental**), podemos observar uma relação como a ilustrada pela linha vermelha na Figura 18.12. A linha de regressão passará pela altura média de todos os genitores e a altura média de toda a prole, que será igual uma à outra porque não ocorreu nenhuma mudança na população entre as gerações. Além disso, os genitores mais altos têm filhos mais altos, e os genitores mais baixos têm filhos mais baixos, e, assim, a inclinação da linha é positiva. Mas a inclinação não é unitária; genitores muito baixos em média têm filhos que são um pouco mais altos, e genitores muito altos em média têm filhos que são um pouco mais baixos que eles mesmos. Essa inclinação de menos que a unidade surge porque a herdabilidade é menos que perfeita. Se o genótipo fosse herdado aditivamente com fidelidade completa, então as alturas da prole seriam idênticas aos valores médios parentais e a linha de regressão seria 1. Por outro lado, se a prole

Um método para avaliar a herdabilidade restrita, h^2

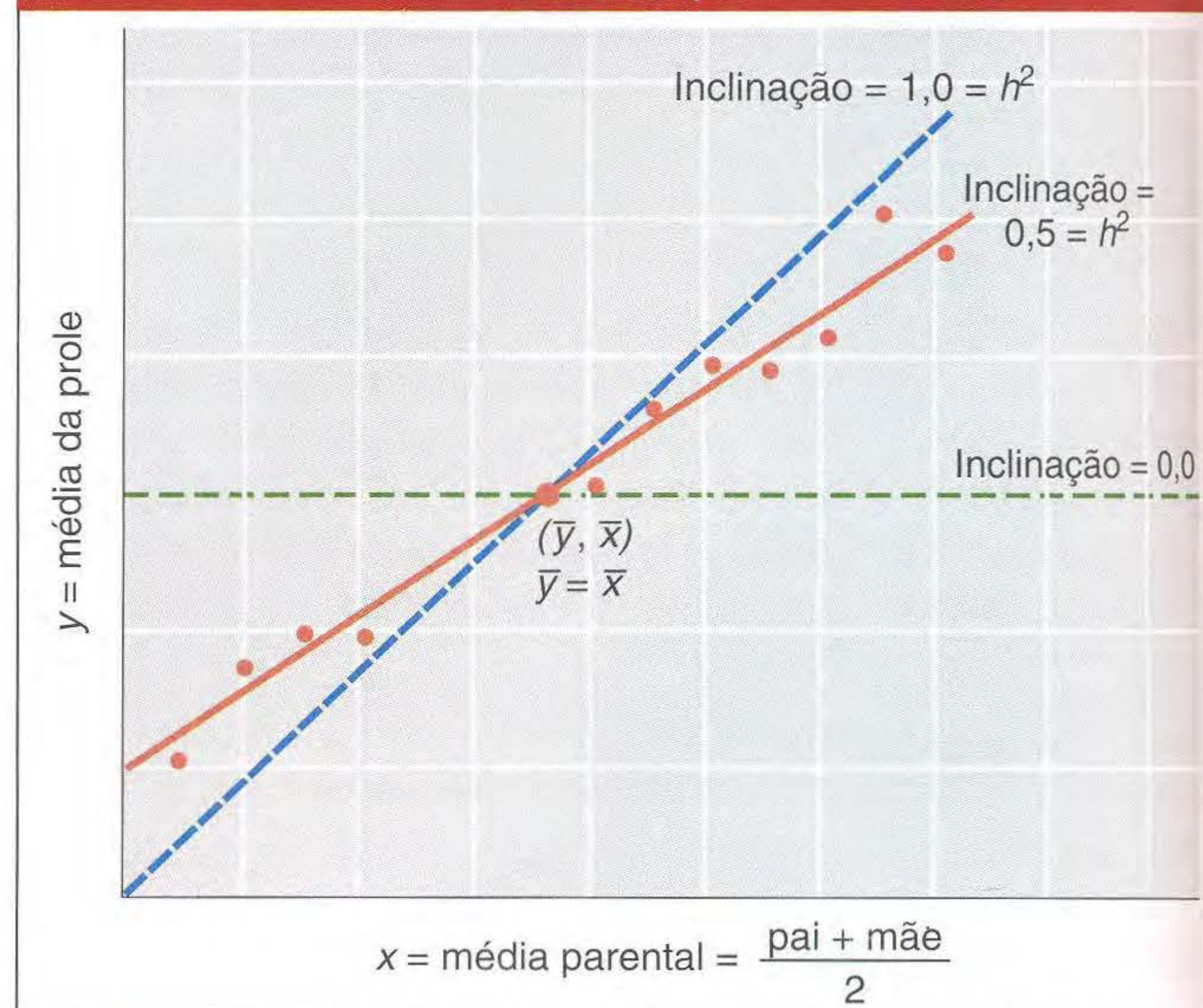


FIG. 18.12 A regressão (linha vermelha) das medidas da prole, y , nos valores parentais médios, x , para uma característica com herdabilidade restrita, h^2 , de 0,5. A linha azul seria a inclinação de regressão se a característica fosse perfeitamente herdável.

não tivesse similaridade de herdabilidade com seus genitores, todos os genitores teriam uma prole da mesma altura média e a inclinação da linha seria 0. Esse raciocínio sugere que a linha de regressão que relaciona o valor da prole com o valor médio parental dá uma estimativa da herdabilidade aditiva. De fato, a inclinação da linha pode ser mostrada matematicamente como sendo uma estimativa correta de h^2 .

O fato de que a inclinação da linha de regressão estima a herdabilidade aditiva nos permite usar h^2 para prever os efeitos da seleção artificial. Suponha que selecionamos pais para a geração seguinte que são, em média, 2 unidades acima da média geral da população da qual foram escolhidos. Se $h^2 = 0,5$, então a prole dos genitores selecionados ficará $0,5(2,0) = 1,0$ unidade acima da média da população parental, porque a inclinação da linha de regressão prevê que um aumento de y resulta de um aumento unitário em x . Podemos definir o **diferencial de seleção** como a diferença entre os genitores selecionados e a média de toda a população em sua geração, e podemos definir a **resposta de seleção** como a diferença entre a prole de genitores selecionados e a média da geração parental. Assim:

$$\text{resposta de seleção} = h^2 \times \text{seleção diferencial}$$

ou

$$h^2 = \frac{\text{resposta de seleção}}{\text{diferencial de seleção}}$$

A segunda expressão nos dá um outro modo de estimar h^2 : fazendo cruzamentos seletivos para uma geração e comparando a resposta de seleção com a seleção diferencial. Geralmente esse processo é feito por várias gerações com o uso da mesma seleção diferencial, e a resposta média é usada como uma estimativa de h^2 .

Lembre que qualquer estimativa de h^2 , como para H^2 , depende da suposição de não haver correlação entre a similaridade dos ambientes dos indivíduos e a similaridade de seus genótipos. Além disso, h^2 em uma população em um conjunto de ambientes não será a mesma que h^2 em uma população diferente em um conjunto diferente de ambientes. Para ilustrar esse princípio, a Figura 18.13 mostra a gama de herdabilidades no sentido restrito relatadas em vários estudos para um número de características em galinhas. Para a maioria das características nas quais foi relatada uma substancial herdabilidade, existem grandes diferenças de estudo para estudo, supostamente porque populações diferentes têm quantidades diferentes de variação genética, e porque foram feitos estudos diferentes em ambientes diferentes. Assim, os criadores que querem saber se a seleção será efetiva em mudar alguma característica em suas galinhas não podem contar com as herdabilidades encontradas em estudos anteriores, mas podem estimar a herdabilidade na população em particular e o ambiente particular no qual o programa de seleção está sendo feito.

Seleção artificial

Um grande registro demonstra a efetividade da seleção artificial em mudar os fenótipos dentro de uma população. Os



FIG. 18.13 Faixas de herdabilidades, h^2 , relatadas para uma variedade de características em galinhas. [De I.M. Lerner and W.J. Libby, *Heredity, Evolution, and Society*. 2nd ed. Copyright 1976 by W.H. Freeman and Company. Photograph copyright by Kenneth Thomas/Photo Researchers.]

cruzamentos de animais e plantas têm, por exemplo, aumentado a produção de leite em vacas e a resistência à ferrugem no trigo. Os experimentos de seleção no laboratório fizeram grandes mudanças na fisiologia e morfologia de muitos organismos, incluindo microrganismos, plantas e animais. Não é possível nenhuma análise desses experimentos em termos de frequências alélicas, pois os loci individuais não foram identificados e seguidos. Entretanto, está claro que as mudanças genéticas têm ocorrido porque as populações mantêm suas características mesmo após a seleção ter terminado. A Figura 18.14 mostra, como um exemplo, que um experimento de seleção obteve grandes mudanças no número médio de cerdas em uma população de *D. melanogaster*. A Figura 18.15 mostra o aumento no número de ovos postos por galinhas como consequência de 30 anos de seleção.

O método usual de seleção para uma característica de variação contínua é a **seleção truncada**. Os indivíduos em determinada geração são reunidos (independentemente de suas famílias), uma amostra é medida, e apenas os indivíduos acima (ou abaixo) de determinado valor fenotípico (o ponto de truncção) são escolhidos como genitores para a geração seguinte.

Uma experiência comum nos programas de seleção artificial é que, à medida que a população se torna mais e mais extrema, sua viabilidade e fertilidade diminuem. Como resultado, nenhum progresso sob seleção é possível, a despeito da presença da variação genética aditiva para a característica, pois os indivíduos selecionados não se reproduzem. A perda de adaptabilidade pode ser um efeito fenotípico direto dos genes para a característica selecionada, em cujo caso nada

A seleção altera o número de cerdas

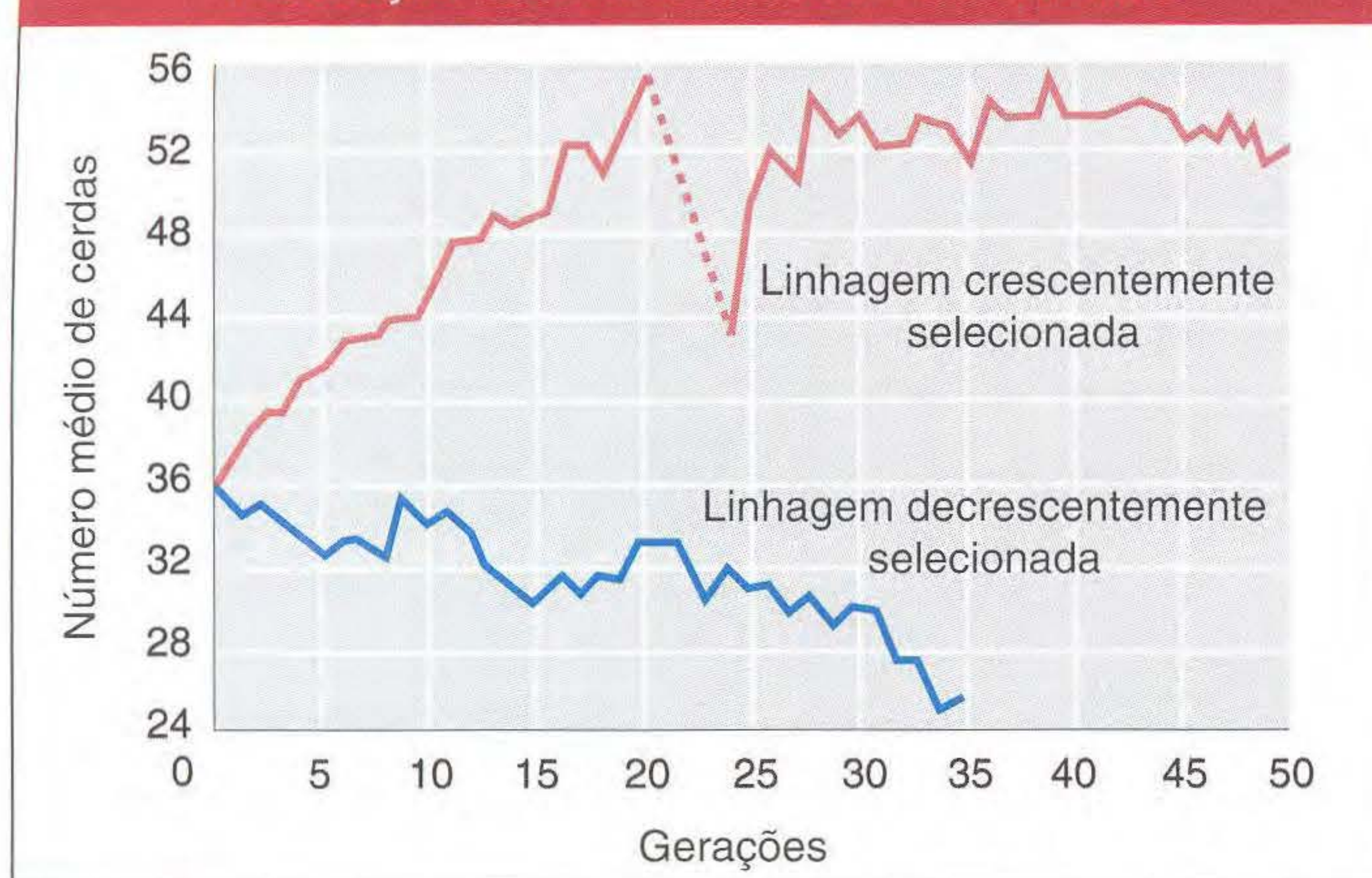


FIG. 18.14 Mudanças no número médio de cerdas obtidas em duas populações de laboratório de *Drosophila melanogaster* por seleção artificial para alto número de cerdas e baixo número de cerdas na outra. A linha tracejada mostra cinco gerações durante as quais não houve prática de seleção. [De K. Mather and B.J. Harrison, "The Manifold Effects of Selection," *Heredity* 3, 1949, 1.]

pode ser feito para melhorar mais a população. Em geral, entretanto, a perda de adaptabilidade está ligada não aos genes que estão sob seleção, mas aos genes de esterilidade que são levados junto com eles. Em tais casos, várias gerações são deixadas cruzar sem seleção até que os recombinantes se formem por acaso, liberando os genes sob seleção de sua associação à esterilidade. A seleção pode então ser continuada, como na linha selecionada superior na Figura 18.14.

Devemos ser muito cuidadosos em nossa interpretação dos programas de seleção natural agrícola em longo prazo. No mundo real da agricultura, as mudanças nos métodos de cultivo, maquinaria, fertilizantes, inseticidas, herbicidas, e assim por diante, estão tomando lugar junto com a produção de variedades geneticamente melhoradas. Os aumentos na produtividade média são conseqüências de todas essas mudanças. Por exemplo, a produtividade média de milho nos EUA melhorou de 40 para 80 alqueires por acre entre 1940 e

A seleção aumenta a produção de ovos



FIG. 18.15 Mudanças na produção média de ovos em uma população de galinhas selecionadas por seu aumento na taxa de pôr ovos em um período de 30 anos. [De I.M. Lerner and W.J. Libby, *Heredity, Evolution, and Society*, 2nd ed. Copyright 1976 by W.H. Freeman and Company. Dados por cortesia de D.C. Lowry.]

1970. Mas os experimentos comparando variedades antigas e novas de milho em ambientes comuns mostraram que apenas cerca de metade desse aumento é o resultado direto de novas variedades de milho (a outra metade sendo o resultado de técnicas agrícolas melhoradas). Além disso, as novas variedades são superiores às antigas apenas em altas densidades de plantio moderno para as quais são selecionadas.

O uso de h^2 em endogamia

Muito embora h^2 seja um número aplicado apenas a determinada população e determinado conjunto de ambientes, ainda é de grande importância para os criadores. Um geneticista de aves interessado em aumentar, digamos, a taxa de crescimento de galinhas não está interessado na variância genética em todas as aves e todas as distribuições ambientais. Para determinado conjunto de aves (ou uma escolha entre alguns conjuntos) nas condições ambientais aproximadamente presentes na prática em fazendas, a pergunta que surge é: Pode um esquema de seleção ser criado para aumentar a taxa de crescimento, e, assim sendo, o quão rápido ela pode ser aumentada? Se um bando tiver uma grande variância genética para a taxa de crescimento e outro tiver apenas um pouco, o criador escolherá o primeiro bando para fazer a seleção. Se a herdabilidade no bando escolhido for muito alta, então a média da população responderá rapidamente à seleção imposta, porque a maioria da superioridade dos genitores selecionados aparecerá na prole. Quanto maior a h^2 , maior a correlação genitor-prole. Se, por outro lado, h^2 é baixa, então apenas uma pequena fração da superioridade dos genitores selecionados aparecerá na geração seguinte.

Se h^2 é muito baixa, pode ser necessário algum esquema alternativo de seleção ou criação. Nesse caso, H^2 juntamente com h^2 pode ser útil ao criador. Suponha que h^2 e H^2 são ambas baixas, o que significa que há uma grande proporção de variância ambiental comparada à variância genética. Algum esquema de redução de s_e^2 deve ser usado. Um método é usar as condições de criação de modo que a variância ambiental seja diminuída. Outra é usar **seleção familiar**. Em vez de selecionar os melhores animais, o criador deixa que os pares produzam várias proles aleatórias, e são selecionados os pares parentais para produzir a geração seguinte com base no desempenho médio dessas proles. A média das proles permite que uma variação ambiental descontrolada e o ruído desenvolvimental sejam cancelados, e pode ser feita uma melhor estimativa da diferença genotípica entre os pares de modo que os melhores pares possam ser escolhidos como genitores da geração seguinte.

Se, por outro lado, h^2 é baixa mas H^2 é alta, então não há muita variância ambiental. A h^2 baixa é o resultado de uma pequena proporção de variância genética aditiva comparada com a variância de dominância. Tal situação pede esquemas especiais de criação que usem variância não-aditiva. Um desses esquemas é o **método híbrido-endogâmico**, que é usado quase universalmente para milho. Um grande número de linhagens endogâmicas são criadas por autofecundação. Essas linhagens endogâmicas são então cruzadas em muitas combinações diferentes (todas as combinações possíveis, se

for economicamente factível), e é escolhido o melhor híbrido. Então são desenvolvidas novas linhagens endogâmicas a partir do melhor híbrido, e novamente são feitos cruzamentos para encontrar o melhor. Esse processo é continuado ciclo após ciclo. Esse esquema seleciona não só os efeitos aditivos, mas também os efeitos para dominância, pois seleciona os melhores heterozigotos como genitores do ciclo seguinte. Ele tem sido a base dos principais avanços genéticos em milho híbrido nos EUA desde 1930. A produção em milho não parece, porém, ter grandes quantidades de variação genética, e, assim, é questionável se essa técnica *ultimamente* gera variedades de mais alta produtividade do que as que resultaram de anos de técnicas de simples seleção com base na variação aditiva.

O método híbrido-endogâmico foi introduzido na criação de todos os tipos de plantas e animais. Tomates e galinhas, por exemplo, hoje são quase exclusivamente híbridos. Também foram feitas tentativas para cruzar trigo híbrido, mas até agora os híbridos obtidos não tiveram uma produtividade melhor do que as variedades não-híbridas hoje usadas.

A subdivisão de variações genética e ambiental dá informações importantes sobre a ação gênica que pode ser usada na criação de plantas e animais.

► 18.8 Localizando os Genes

Não é possível identificar todos os genes que influenciam o desenvolvimento de determinada característica usando técnicas puramente genéticas. Em uma dada população, apenas um subgrupo de genes que contribuem para o desenvolvimento de qualquer determinada característica será geneticamente variável. Assim, apenas parte da variação possível será observada. Isso é verdade mesmo para genes que determinam simples características qualitativas, como, por exemplo, os genes que determinam a configuração antigênica total da membrana das hemácias humanas. Cerca de 40 loci determinando os grupos sanguíneos humanos são atualmente conhecidos. Cada um foi descoberto achando pelo menos uma pessoa com uma especificidade imunológica que difere das especificidades de outras pessoas. Muitos outros loci que determinam a estrutura da membrana das hemácias podem não ter sido ainda descobertos porque todas as pessoas estudadas são geneticamente idênticas. A análise *genética* detecta genes apenas quando há alguma variação alélica. Em contraste, a análise *molecular* lida diretamente com o DNA e sua informação traduzida, e, assim, pode identificar genes como trechos de DNA que codificam alguns produtos, mesmo quando eles não variam, desde que os produtos gênicos possam ser identificados.

Muito embora uma característica possa apresentar variação fenotípica contínua, a base genética para as diferenças pode ser a variação alélica em um único locus. A maior parte das mutações clássicas em *Drosophila* são fenotipicamente variáveis em sua expressão, e, em muitos casos, a classe mutante difere pouco do tipo selvagem, e, assim, muitas mos-

cas que possuem a mutação são indistinguíveis das moscas normais. Mesmo os genes do complexo *bithorax*, que têm marcantes mutações homeóticas que transformam os halteres em asas (veja Figura 12.1), também têm alelos fracos que aumentam o tamanho dos halteres apenas ligeiramente em média, e, assim, as moscas do genótipo mutante podem parecer do tipo selvagem.

Às vezes é possível usar um conhecimento anterior de bioquímica e desenvolvimento de um organismo para supor que a variação em um locus conhecido é responsável por, pelo menos, alguma variação em determinado caráter. Tal locus é um **gene candidato** para a investigação de variação fenotípica contínua. A variação na atividade da enzima fosfatase ácida nas hemácias foi investigada desse modo. Como estamos lidando com variação na atividade enzimática, uma boa hipótese seria admitir que há uma variação alélica no locus que codifica essa enzima. Quando Harry Harris e David Hopkinson colheram uma amostra da população inglesa, eles descobriram que havia, de fato, três formas alélicas, A, B e C, que resultavam em enzimas com diferentes níveis de atividade. O Quadro 18.2 mostra a atividade média, a variação em atividade e a frequência populacional dos seis genótipos. A Figura 18.16 mostra a distribuição de atividade de toda a população e como ela é composta de distribuições de genótipos diferentes. Como mostra o Quadro 18.2, da variação em atividade na distribuição total (607,8), cerca de metade é explicada pela variação média dentro dos genótipos (310,7); de modo que metade ($607,8 - 310,7 = 297,1$) deve ser atribuída à variação entre as médias dos seis genótipos. Embora muito da variação em atividade seja explicado pelas diferenças médias entre os genótipos, resta a variação dentro de cada genótipo, que pode ser o resultado de influências ambientais ou de segregação de outros genes ainda não identificados.

A distribuição de uma característica é composta das distribuições de genótipos diferentes

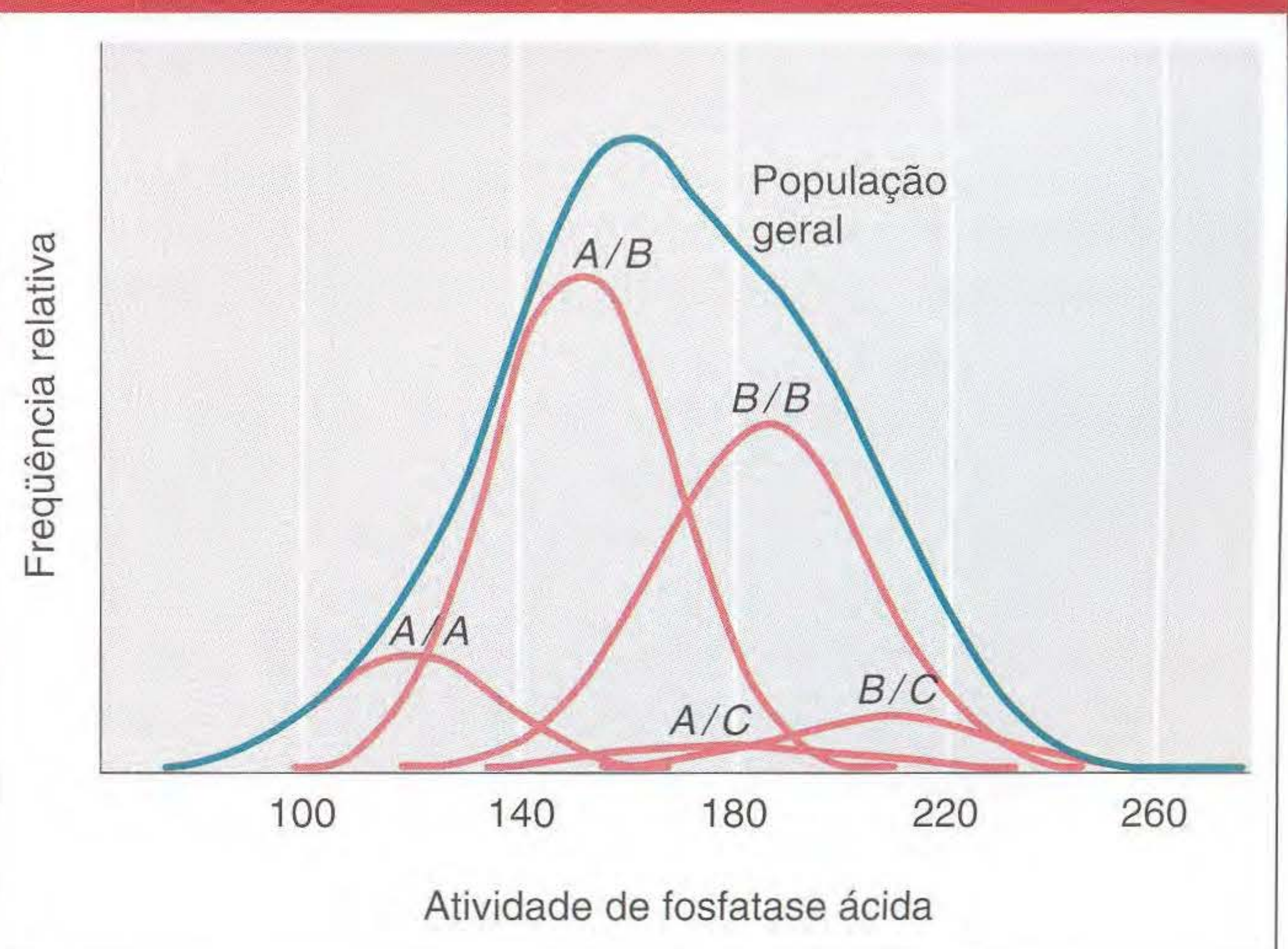


FIG. 18.16 As curvas vermelhas mostram a distribuição de atividade de fosfatase ácida nas hemácias para genótipos diferentes, e a curva verde mostra a distribuição de atividade em uma população inglesa constituída de uma mistura desses genótipos. [H. Harris, *The Principles of Human Biochemical Genetics*, 3rd ed. Copyright 1970 by North-Holland.]

Quadro 18.2 Atividade das Hemácias de Genótipos Diferentes de Fosfatase Ácida na População Inglesa

Genótipo	Atividade média	Variança de atividade	Frequência na população
A/A	122,4	282,4	0,13
A/B	153,9	229,3	0,43
B/B	188,3	380,3	0,36
A/C	183,8	392,0	0,03
B/C	212,3	533,6	0,05
C/C	240	—	0,002
Média geral	166,0	310,7	
Distribuição total		607,8	

Nota: As médias são avaliadas pela frequência na população.

Fonte: H. Harris, *The Principles of Human Biochemical Genetics*, 3rd ed. North-Holland, 1980.

Usando o método de gene candidato, geralmente observamos que parte da variação em uma população é atribuível a alelos diferentes em um único locus, mas a proporção da variança associada ao único locus geralmente é menor do que a que foi encontrada para a atividade da fosfatase ácida. Por exemplo, os três alelos comuns para o gene *apoE*, que codificam a proteína apolipoproteína E, contribuem apenas com 16% da variança nos níveis sanguíneos das lipoproteínas de baixa densidade que levam colesterol e estão implicadas no excesso dos níveis de colesterol. A variança restante é uma consequência de alguma combinação desconhecida de variação genética em outros loci e variação ambiental.

Às vezes, os genes candidatos para uma variação fenotípica em uma espécie podem ser deduzidos de estudos experimentais de outras espécies. Esse método foi usado para localizar um gene que contribui para as diferenças de cor da pele entre os africanos e europeus. A pigmentação normal do peixe-zebra é uma consequência da deposição de grânulos de melanina nas células da pele. O peixe-zebra com uma mutação, *golden*, tem um pigmento muito mais claro do que o normal porque a mutação altera a sequência de aminoácidos de uma proteína que participa da deposição de melanina. A genômica comparativa mostrou que o gene que codifica essa proteína foi conservado nos vertebrados, incluindo os humanos, e que a introdução de mRNA produzido pela versão humana do gene irá restaurar a formação de pigmento no peixe-zebra *golden* mutante. Quando populações humanas foram amostradas, as populações européias eram caracterizadas por um alelo variante desse gene que resulta em deposição deficiente de pigmento, enquanto as populações africanas tinham o alelo ancestral. Um estudo examinou a relação entre o estado alélico e a cor da pele em indivíduos da população afro-americana mista ancestral. Há uma diferença de cerca de 30 unidades de melanina entre a média da população mista e a da população européia, e o estudo descobriu que, dessa diferença de 30 unidades, 9,4 unidades (31%) podem ser atribuídas a diferenças alélicas para esse

gene. Embora esse efeito seja grande, cerca de 70% da variação contínua em cor da pele em toda a espécie humana ainda não foi explicada, e o número de genes envolvidos é desconhecido. Como a variação em cor da pele é quase totalmente genética, é uma característica que é muito mais favorável para uma análise locus por locus do que o para forma ou tamanho, cujo desenvolvimento é fortemente influenciado tanto por genes quanto pelo ambiente, para as quais a variedade dos efeitos gênicos no crescimento nos leva a esperar que um grande número de genes tenha alguma influência.

Segregação de genes marcadores

Os genes que segregam para uma característica quantitativa, chamados de **loci de características quantitativas**, ou **QTL**, não podem ser identificados na maioria dos casos. É possível, entretanto, localizar regiões do genoma nas quais estão os loci relevantes e estimar quanto da variação total é atribuída pela variação QTL em cada região. Essa análise pode ser feita em organismos experimentais cruzando duas linhagens que diferem acentuadamente na característica quantitativa, e que também diferem em alelos em loci bem conhecidos, chamados **genes marcadores**. Os genes marcadores usados para tais análises são genes para os quais genótipos diferentes podem ser distintos por alguns fenótipos visíveis que não podem ser confundidos com características quantitativas (por exemplo, cor de olho em *Drosophila*) ou pela mobilidade eletroforética das proteínas que eles codificam ou pela sequência de DNA dos próprios genes. Um experimento típico envolve cruzar duas linhagens que diferem marcadamente na característica quantitativa e, também, em alelos marcadores. A F_1 resultante do cruzamento entre as duas linhagens pode então ser cruzada entre si para fazer uma F_2 segregante ou pode ser retrocruzada com uma das linhagens parentais. Se existirem QTL proximamente ligados a um gene marcador, então os genótipos marcadores diferentes e os QTL serão herdados juntos, e os genótipos marcadores diferentes na F_2 , ou retrocruzados, terão fenótipos médios diferentes para a característica quantitativa.

Análise de ligação quantitativa

A localização de QTL em pequenas regiões dentro de cromossomos requer a presença de loci marcadores espaçados ao longo do cromossomo. Além disso, deve ser possível criar linhagens parentais que diferem uma da outra nos alelos levados por esses loci marcadores. Com o advento de técnicas moleculares que podem detectar polimorfismos genéticos a nível de DNA, foi descoberta uma densidade muito alta de loci variantes em cromossomos de todas as espécies. São particularmente úteis os polimorfismos de comprimento de fragmento de restrição (RFLP), as repetições em tandem e os polimorfismos de um só nucleotídeo (SNP) no DNA. Tais polimorfismos são tão comuns que quaisquer duas linhagens selecionadas para uma diferença em características quantitativas também diferem uma da outra em alguns loci marcadores moleculares conhecidos, espaçados algumas unidades de crossing uns dos outros em cada cromossomo.

Um protocolo experimental para localizar genes, mostrado na Figura 18.17, usa grupos de indivíduos que diferem marcadamente na característica de interesse bem como em loci marcadores. Esses grupos podem ser criados por várias gerações de seleção divergente para criar linhagens extremas, ou pode ser obtida uma vantagem das variedades existentes ou dos grupos familiares que diferem marcadamente nessa característica. Essas linhagens devem então ser examinadas quanto a loci marcadores que difiram entre si. É feito um cruzamento entre as duas linhagens, e a F_1 é então cruzada para produzir uma F_2 segregante ou é cruzada com uma das linhagens parentais para produzir um retrocruzamento

segregante. Um grande número de prole da geração segregante é então avaliado quanto ao fenótipo quantitativo e caracterizado por seu genótipo nos loci marcadores. Um locus marcador que é improvável ou pouco proximamente ligado a qualquer QTL que afete uma característica de interesse terá o mesmo valor médio de características quantitativas para todos os seus genótipos, enquanto um que é proximamente ligado a alguns QTL irá diferir em seu fenótipo quantitativo médio de um genótipo marcador para outro.

Quanta diferença há no fenótipo quantitativo médio entre genótipos marcadores depende tanto da intensidade do efeito de QTL quanto da intensidade de ligação entre QTL e o locus marcador. Suponha, por exemplo, que duas linhagens selecionadas difiram por um total de 100 unidades em alguma característica quantitativa. A linhagem com o valor mais alto é homozigota $+/+$ em determinado QTL, enquanto a linhagem com o valor mais baixo é homozigota $-/-$, e cada alelo $+$ nesses QTL contribui com 5 unidades da diferença total entre as linhagens. Suponha, ainda, que a linhagem alta seja M/M e a linhagem baixa m/m em um locus marcador distante 10 unidades de crossing de QTL. Então, como mostrado na Figura 18.17, existem 4 unidades de diferença entre o gameta médio levando o alelo M e o gameta médio levando um alelo m na F_2 segregante. Podemos, portanto, calcular que 8 unidades de diferença entre o homozigoto M/M e um homozigoto m/m são atribuíveis a esse QTL. Assim, temos 8% de diferença média entre as linhagens selecionadas originais. O QTL de fato contribui com 10% da diferença. A discrepância vem da recombinação entre o gene marcador e QTL. Podemos então repetir esse processo usando loci marcadores em outros locais ao longo do cromossomo e em cromossomos diferentes para avaliar outras frações da diferença quantitativa entre as linhagens selecionadas originais.

Essa técnica foi usada para situar segmentos cromossômicos associados a características tais como peso dos frutos em tomates, número de cerdas em *Drosophila* e características vegetativas em milho. No caso do milho, 82 características vegetativas foram examinadas em um cruzamento entre linhagens que diferem em 20 marcadores de DNA. Em média, cada característica estava significativamente associada a 14 marcadores diferentes, mas a proporção da diferença de características entre as duas linhagens que estavam associadas a qualquer marcador em particular era geralmente muito pequena. A Figura 18.18 mostra a proporção das associações estatisticamente significativas entre marcador e característica (no eixo y) que contribuem para proporções divergentes das diferenças de características entre as linhagens. Como mostra a Figura 18.18, a maioria das associações contribui com menos de 1% da diferença da característica. Infelizmente, em genética humana, embora a segregação de gene marcador possa ser usada para localizar distúrbios monogênicos, o pequeno tamanho dos heredogramas humanos torna a técnica de segregação do marcador inaplicável a loci de características quantitativas, porque a prole é muito pequena para que qualquer determinado marcador forneça alguma precisão.

Um protocolo experimental para localizar genes

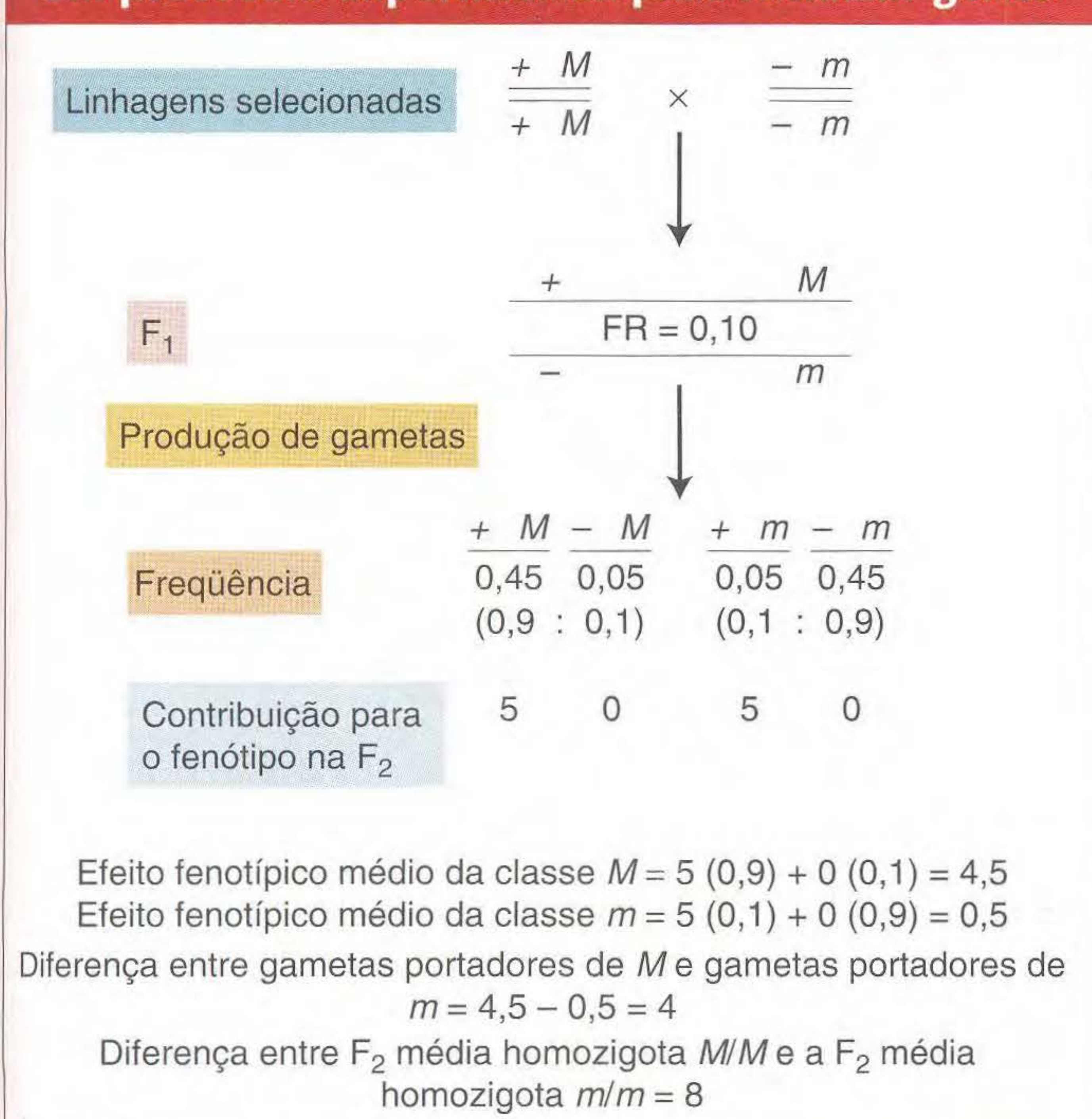


FIG. 18.17 Resultados de um cruzamento entre duas linhagens selecionadas que diferem em QTL e em um locus marcador molecular distante 10 unidades de crossing do QTL. O alelo QTL+ adiciona 5 unidades de diferença ao fenótipo.

Apenas uma pequena percentagem da diferença de características está associada a qualquer marcador de DNA

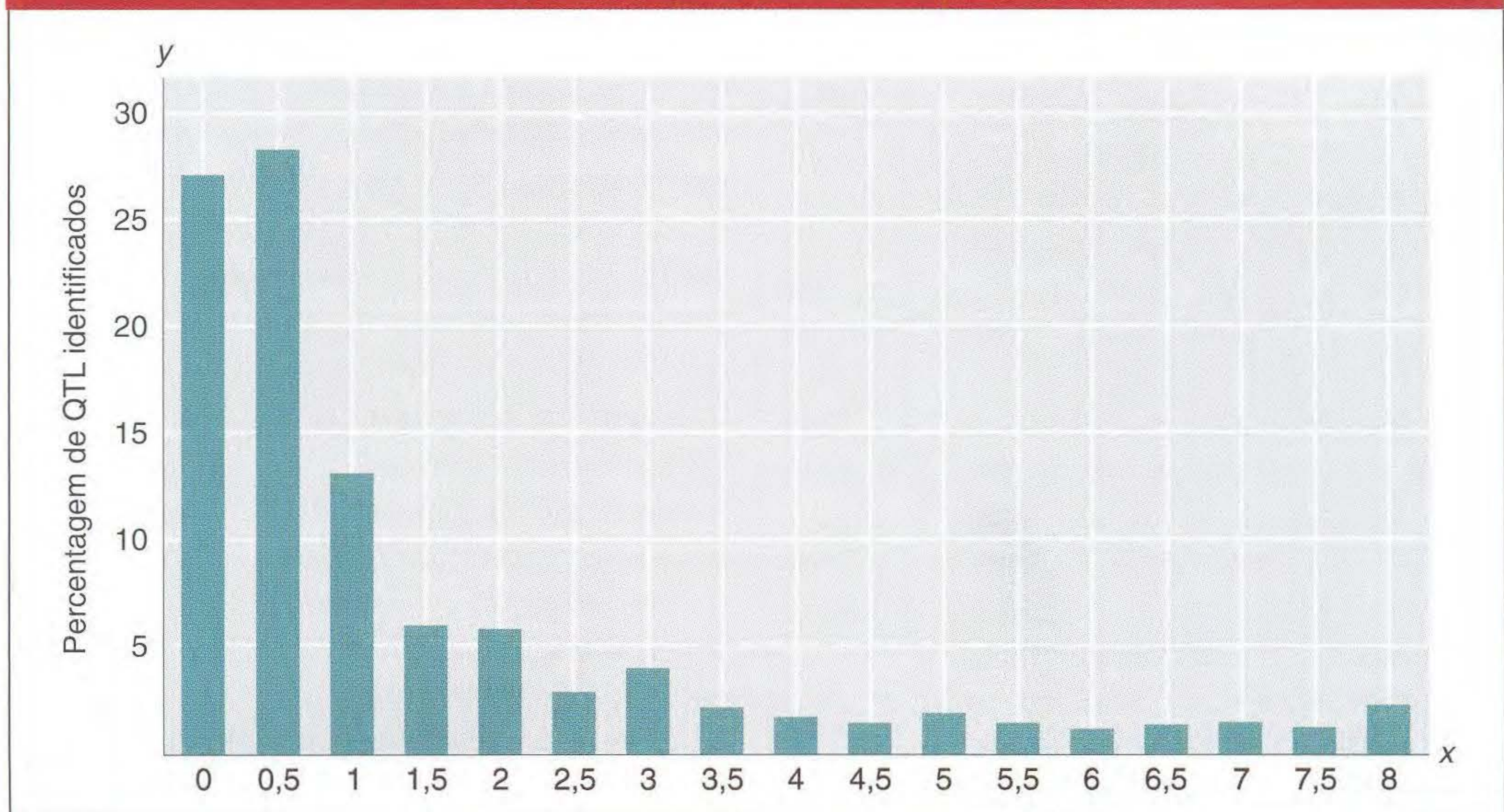


FIG. 18.18 Distribuição de associações das diferenças de características entre duas linhagens de milho com um arranjo de marcadores de DNA. O eixo x mostra a percentagem da diferença explicada entre as duas linhagens em determinada característica que pode ser associada a qualquer gene marcador. O eixo y mostra a proporção de todos os QTL identificados que têm a percentagem correspondente de sua diferença explicada. Note que 55% de todas as associações (primeiras duas colunas) correspondem a menos de 1% de suas diferenças de característica. [De M. Lynch and B. Walsh, *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, 1998. Dados de M.D. Edwards, C.W. Stuber, and J.F. Wendel, *Genetics* 116, 1987, 113-125.]

Para muitos organismos (por exemplo, humanos), não é possível fazer linhagens homozigotas diferindo em alguma característica e, então, cruzá-las para produzir uma geração segregante. Para tais organismos, podemos usar as diferenças entre irmãos levando diferentes alelos marcadores de genitores heterozigotos. Esse método tem muito menos poder de encontrar QTL, sobretudo quando o número de irmãos em qualquer família é pequeno, como ocorre em famílias humanas. Como consequência, as tentativas em mapear QTL para características humanas não têm sido bem-sucedidas, embora a técnica de segregação de marcadores tenha sido um sucesso em encontrar loci cujas mutações são responsáveis por distúrbios monogênicos ou por características quantitativas cuja variação é fortemente influenciada pela

variação em um locus. Por exemplo, as pessoas variam em sua habilidade em sentir o gosto da substância feniltiocarbamida (PTC). Algumas podem detectar concentrações bem baixas, enquanto outras só detectam concentrações altas ou são incapazes de sentir o gosto da PTC. A análise de ligação usando polimorfismos de um único nucleotídeo localizou uma região do cromossomo humano 7q que contribuía com cerca de 75% da variação na sensibilidade gustativa. Essa região cromossômica já era conhecida como contendo vários genes que codificam proteínas receptoras de gosto amargo. Quando o DNA desses genes foi seqüenciado, foram encontrados três polimorfismos de aminoácidos em um dos genes como estando fortemente associados à diferença entre os fenótipos sensível e insensível.

► Apêndice Estatístico

A informação completa sobre a distribuição de um fenótipo em uma população pode ser dada apenas especificando a frequência de cada classe medida, mas uma grande quantidade de informações pode ser resumida apenas em duas estatísticas. Primeiro, precisamos de alguma medida da localização da distribuição ao longo do eixo da medida. (Por exemplo, as medidas individuais da altura de alunos masculinos tendem a se aglomerar ao redor de 100 cm ou 200 cm?) Segundo, precisamos de alguma medida da quantidade de variação dentro da distribuição. (Por exemplo, estão as alturas dos alunos todas concentradas ao redor da medida central ou elas variam bastante em uma grande faixa?)

Medidas da tendência central

A moda A maioria das distribuições dos fenótipos parecem um pouco com as da Figura 18.2: uma única moda está situada perto da metade da distribuição, com frequências diminuindo em ambos os lados. Entretanto, existem exceções a esse padrão. A Figura 18.19a mostra a distribuição muito assimétrica dos pesos das sementes na planta *Crinum longifolium*. A Figura 18.19b mostra uma **distribuição bimodal** (duas modas) de probabilidade de sobrevivência larvar para diferentes homozigotos de segundo cromossomo em *Drosophila willistoni*.

Algumas funções de distribuição são assimétricas ou bimodais

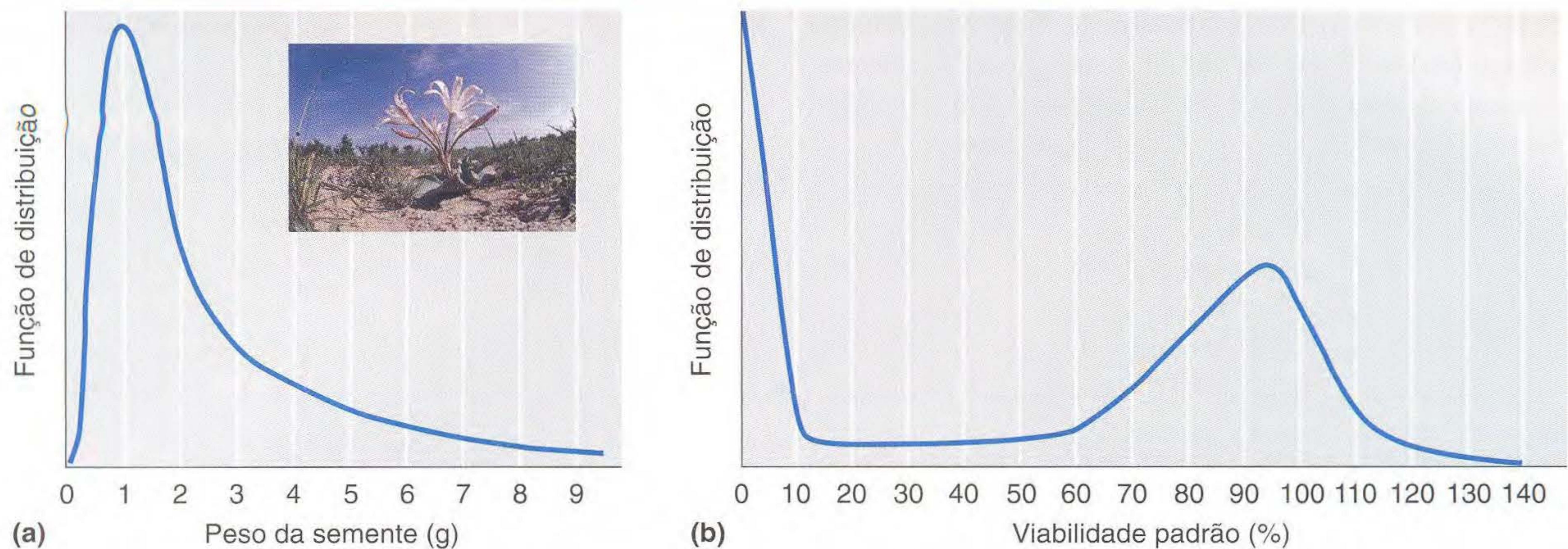


FIG. 18.19 (a) Distribuição assimétrica do peso de sementes em *Crinum longifolium*; (b) distribuição bimodal de sobrevivência de *Drosophila willistoni* expressa como uma porcentagem da sobrevivência padrão. [De S. Wright, *Evolution and the Genetics of Populations*, vol. 1. Copyright 1968 by University of Chicago Press. Foto: Earth Scenes/Copyright Thompson GOSF.]

Uma distribuição bimodal pode indicar que a população em estudo poderia ser mais bem considerada uma mistura de duas populações, cada uma com sua própria moda. Na Figura 18.19b, a moda à esquerda provavelmente representa uma subpopulação de graves mutações de um só locus que são extremamente deletérias quando homozigotas, mas cujos efeitos não são sentidos no estado heterozigoto no qual geralmente existem nas populações naturais. A moda da direita é parte da distribuição dos modificadores de pequeno efeito da viabilidade “normal” da distribuição.

A média Uma medida mais comum da tendência central é a média aritmética, ou a **média**. A média da medida, \bar{x} , é simplesmente a soma de todas as medidas individuais, x_i , dividida pelo número de medidas na amostra, N :

$$\text{média} = \bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{1}{N} \sum x_i$$

onde \sum representa a operação de soma de todos os valores de i de 1 a N , e x_i é a i^{a} medida.

Em uma típica amostra grande, o mesmo valor medido aparecerá mais de uma vez, pois vários indivíduos terão o mesmo valor dentro da precisão do instrumento de medida. Em tal caso, \bar{x} pode ser reescrito como a soma de todos os valores medidos, cada um avaliado pela frequência com que ele aparece na população. De um total de N indivíduos medidos, suponha que n_1 caia em uma classe com valor x_1 , que n_2 caia na classe com valor x_2 , e assim por diante, de modo que $\sum n_i = N$. Se f_i for a **frequência relativa** da i^{a} classe de medida, de modo que

$$f_i = \frac{n_i}{N},$$

então poderemos reescrever a média como

$$\bar{x} = f_1x_1 + f_2x_2 + \dots + f_kx_k = \sum f_ix_i$$

onde x_i é igual ao valor da i^{a} classe medida.

Quadro 18.3 Número de Indentações no Pente Sexual das Pernas Direita (x) e Esquerda (y) e a Soma dos Dois (T) para 20 Machos de *Drosophila*

x	y	T	n_i	$f_i = n_i/N$
6	5	11	1	$\frac{1}{20} = 0,05$
6	6	12	3	$\frac{3}{20} = 0,15$
5	7	12		
6	6	12		
7	6	13	4	$\frac{4}{20} = 0,20$
5	8	13		
6	7	13		
7	6	13		
8	6	14	7	$\frac{7}{20} = 0,35$
6	8	14		
7	7	14		
7	7	14		
7	7	14		
6	8	14		
8	6	14		
8	7	15	3	$\frac{3}{20} = 0,15$
7	8	15		
6	9	15		
8	8	16	2	$\frac{2}{20} = 0,10$
7	9	16		
$N = 20$		$s_x^2 = 0,8275$		$s_x = 0,9096$
$\bar{x} = 6,25$		$s_y^2 = 1,1475$		$s_y = 1,0722$
$\bar{y} = 7,05$		$s_T^2 = 1,71$		$s_T = 1,308$
$\bar{T} = 13,70$		$\text{cov } xy = -0,1325$		
		$r_{xy} = -0,1360$		

Apliquemos esses métodos de cálculo aos dados do Quadro 18.3, que nos dá os números de cerdas tipo dente nos pentes sexuais das pernas dianteiras direita (x) e esquerda (y) e em ambas as pernas ($T = x + y$) de 20 *Drosophila*. Observada no momento apenas a soma das duas pernas, T , encontramos o número médio de pentes sexuais, \bar{T} como sendo

$$\begin{aligned}\bar{T} &= \frac{11 + 12 + 12 + 12 + 13 + \dots + 15 + 16 + 16}{20} \\ &= \frac{274}{20} \\ &= 13,7\end{aligned}$$

Alternativamente, usando as frequências relativas de valores de medidas diferentes, vemos que

$$\begin{aligned}\bar{T} &= 0,05(11) + 0,15(12) + 0,20(13) + 0,35(14) \\ &\quad + 0,15(15) + 0,10(16) \\ &= 13,7\end{aligned}$$

Medidas de dispersão: a variância

Uma segunda característica de uma distribuição é a largura de seu espalhamento ao redor da classe central. Duas distribuições com a mesma média podem diferir muito na proximidade com que as medidas estão concentradas ao redor da média. A medida mais comum da variação ao redor do centro é a **variância**, que é definida como o desvio quadrado médio das observações da média, ou

$$\begin{aligned}\text{variância} &= s^2 \\ &= \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N} \\ &= \frac{1}{N} \sum (x_i - \bar{x})^2\end{aligned}$$

Quando mais de um indivíduo tem o mesmo valor medido, a variância pode ser escrita como

$$\begin{aligned}s^2 &= f_1(x_1 - \bar{x})^2 + f_2(x_2 - \bar{x})^2 + \dots + f_k(x_k - \bar{x})^2 \\ &= \sum f_i(x_i - \bar{x})^2\end{aligned}$$

Para evitarmos subtrair cada valor de x separadamente da média, podemos usar uma fórmula computada alternativa que é algebricamente idêntica à primeira forma da equação:

$$s^2 = \left(\frac{1}{N} \sum x_i^2 \right) - (\bar{x})^2$$

Como a variância está em unidades quadradas (centímetros quadrados, por exemplo), é comum tirar a raiz quadrada da variância, que então tem as mesmas unidades que a medida em si. Essa medida da raiz quadrada da variação é chamada de **desvio padrão** da distribuição:

$$\text{desvio padrão} = s = \sqrt{\text{variância}} = \sqrt{s^2}$$

Os dados para os pentes sexuais no Quadro 18.3 podem ser usados para exemplificar estes cálculos:

$$\begin{aligned}s_T^2 &= \frac{(11 - 13,7)^2 + (12 - 13,7)^2 + (12 - 13,7)^2}{20} \\ &\quad + \dots + \frac{(15 - 13,7)^2 + (16 - 13,7)^2}{20} \\ &= \frac{34,20}{20} = 1,71\end{aligned}$$

Também podemos usar a fórmula computada que evita obter desvios individuais:

$$s_T^2 = \frac{1}{N} \sum T_i^2 - \bar{T}^2 = \frac{3.788}{20} - 187,69 = 1,71$$

e

$$s = \sqrt{1,71} = 1,308$$

A Figura 18.20 mostra duas distribuições tendo a mesma média mas desvios padrões diferentes (curvas A e B) e duas distribuições tendo o mesmo desvio padrão mas médias diferentes (curvas A e C).

A média e a variância de uma distribuição não a descrevem completamente. Elas não distinguem uma distribuição simétrica de uma assimétrica, por exemplo. Existem sete distribuições simétricas que têm a mesma média e variância mas ainda têm formas um pouco diferentes. Entretanto, para lidar com a maioria dos problemas de genética quantitativa, a média e a variância são suficientes para caracterizar uma distribuição.

Medidas de relação

Covariância e correlação Outra noção estatística que é de uso no estudo de genética quantitativa é a associação, ou **correlação**, entre variáveis. Como resultado de vias complexas de causa, muitas variáveis na natureza variam juntas, mas de modo imperfeito ou aproximado. A Figura 18.21a fornece um exemplo, mostrando os tamanhos de dois determinados dentes em várias espécies animais de um mamífero fóssil,

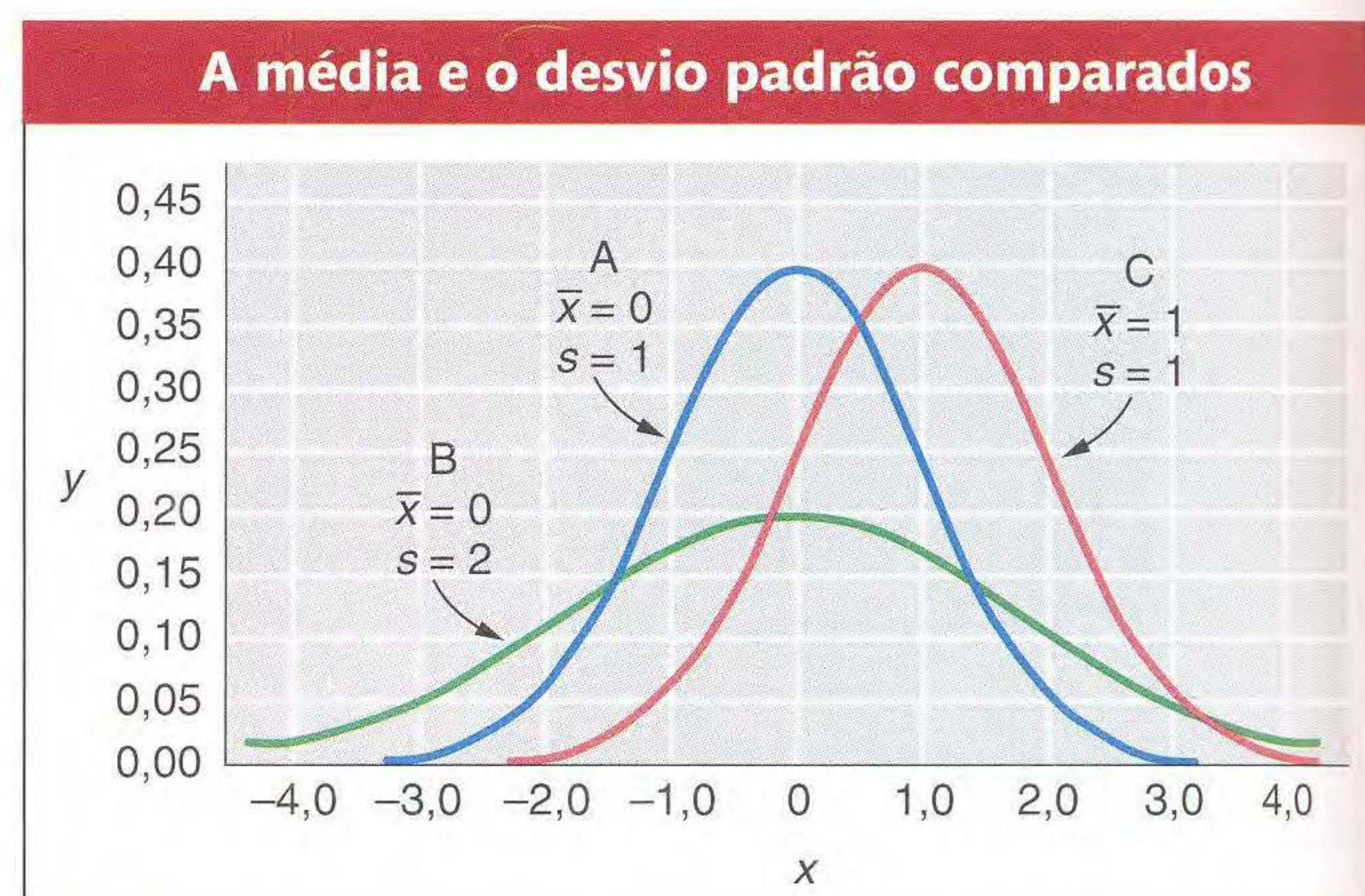


FIG. 18.20 Três funções de distribuição, duas das quais têm a mesma média (A e B) e duas das quais têm o mesmo desvio padrão (B e C).

Exemplos de correlação imperfeita (a) e correlação próxima (b)

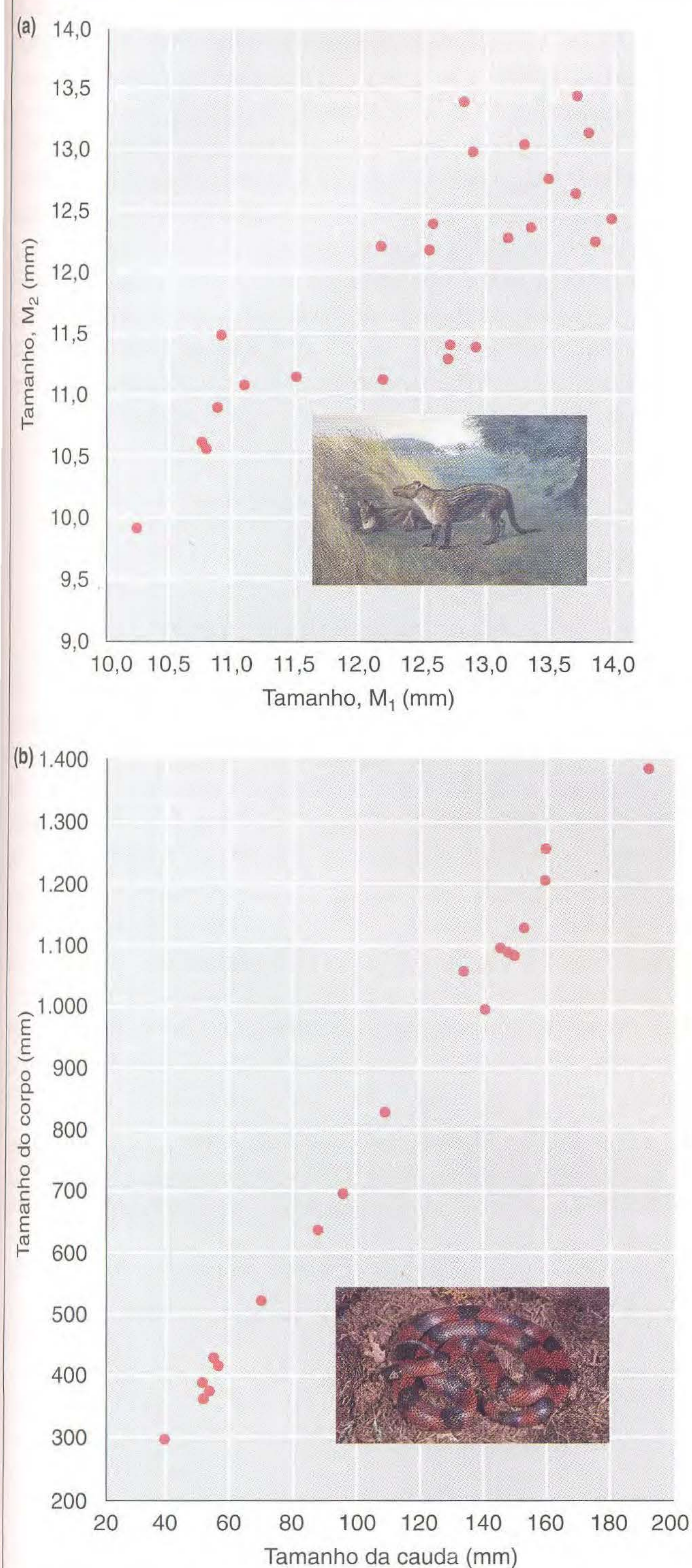


FIG. 18.21 Diagrama disperso das relações entre pares de variáveis. (a) Relação entre os tamanhos do primeiro e segundo molares inferiores (M_1 e M_2) no mamífero extinto *Phenacodus primaevus*. Cada ponto dá as medidas M_1 e M_2 de um indivíduo. (b) Tamanho da cauda e do corpo de 18 cobras *Lampropeltis polyzona*. [(a) Imagem: Negative nº 2430, fotografia de Charles Knight; cortesia do Department of Library Services, American Museum of Natural History. (b) Foto: Animals Animals/Copyright Zig Leszczynski.]

Phenacodus primaevus. Há uma ligeira tendência de que indivíduos com primeiros molares maiores tendam a ter segundos molares maiores, mas há uma considerável dispersão de pontos de dados ao redor dessa tendência. Em contraste, a Figura 18.21b mostra que o tamanho do corpo e o tamanho da cauda de cobras (*Lampropeltis polyzona*) são bem relacionados um ao outro, com todos os pontos caindo próximo a uma linha reta, que pode ser traçada entre eles da parte inferior esquerda do gráfico para a superior direita.

A medida usual da precisão de uma relação entre duas variáveis x e y é o **coeficiente de correlação**, r_{xy} . Ele é calculado em parte do produto da derivação de cada observação de x da média dos valores de x e o desvio de cada observação de y da média de valores y , uma quantidade chamada de **covariança** de x e y ($\text{cov } xy$):

$$\begin{aligned}\text{cov } xy &= \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \dots}{N} \\ &\quad + \frac{(x_N - \bar{x})(y_N - \bar{y})}{N} \\ &= \frac{1}{N} \sum (x_i - \bar{x})(y_i - \bar{y})\end{aligned}$$

Uma fórmula que é algebricamente exatamente equivalente mas que torna a computação mais fácil é

$$\text{cov } xy = \left(\frac{1}{N} \sum x_i y_i \right) - \bar{x}\bar{y}$$

Usando essa fórmula, podemos calcular a covariança entre as contagens direita (x) e esquerda (y) no Quadro 18.3.

$$\begin{aligned}\text{cov } xy &= \left(\frac{1}{N} \sum xy \right) - \bar{x}\bar{y} \\ &= \frac{(6)(5) + (6)(6) + \dots + (8)(8) + (7)(9)}{20} \\ &\quad - (6,65)(7,05) \\ &= -0,1325\end{aligned}$$

A correlação, r_{xy} , é definida como

$$\text{correlação} = r_{xy} = \frac{\text{cov } xy}{s_x s_y},$$

Na fórmula para correlação, os produtos dos desvios são divididos pelo produto dos desvios padrões de x e y (s_x e s_y). Essa normalização dos desvios padrões têm o efeito de fazer r_{xy} um número menos dimensional que é independente das unidades nas quais x e y são medidos. Assim definido, r_{xy} irá variar de -1 , que significa uma relação negativa perfeitamente linear entre x e y de $+1$, o que indica uma relação positiva perfeitamente linear entre x e y . Se $r_{xy} = 0$, não há relação linear entre as variáveis. Os valores intermediários a 0 e -1 ou $+1$ indicam graus intermediários de relação entre as variáveis. Os dados na Figura 18.21a e 21b têm valores r_{xy} de $0,82$ e $0,99$, respectivamente. No exemplo dos pentes

sexuais do Quadro 18.3, a correlação entre as pernas esquerda e direita é

$$r_{xy} = \frac{\text{cov } xy}{\sqrt{s_x^2 s_y^2}} = \frac{-0,1325}{\sqrt{(0,8275)(1,1475)}} = 0,1360$$

um valor muito pequeno. É importante notar, entretanto, que, às vezes, quando não há relação *linear* entre duas variáveis, mas há uma relação *não-linear* regular entre elas, uma variável pode ser perfeitamente prevista pela outra. Considere, por exemplo, a parábola mostrada na Figura 18.22. Os valores de y são perfeitamente previsíveis pelos valores de x ; mas $r_{xy} = 0$, porque, na média de toda a faixa de valores x , valores maiores de x estão associados a valores maiores ou menores de y .

Correlação e igualdade É importante notar que a correlação entre dois grupos de números não é o mesmo que identidade numérica. Por exemplo, dois conjuntos de valores podem ser perfeitamente *correlatos*, muito embora os valores em um conjunto sejam muito maiores que os valores no outro conjunto. Considere os seguintes pares de valores:

x	y
1	22
2	24
3	26

As variáveis x e y nos pares são perfeitamente correlacionadas ($r = 1,0$), embora cada valor de y seja cerca de 20 unidades maior que o valor correspondente de x . Duas variáveis são perfeitamente correlacionadas se, para um aumento unitário de um, há um aumento constante no outro (ou uma diminuição constante se r for negativo).

A importância da diferença entre correlação e identidade surge quando consideramos o efeito do ambiente nas características herdáveis. Os genitores e a prole podem estar perfeitamente correlacionados em alguma característica tal como a altura, mas, devido a uma diferença ambiental entre gerações, cada criança pode ser mais alta que seus genitores. Esse fenômeno aparece nos estudos de adoção, nos quais as crianças podem estar correlacionadas a seus pais biológicos,

Uma parábola é um exemplo de uma relação não-linear

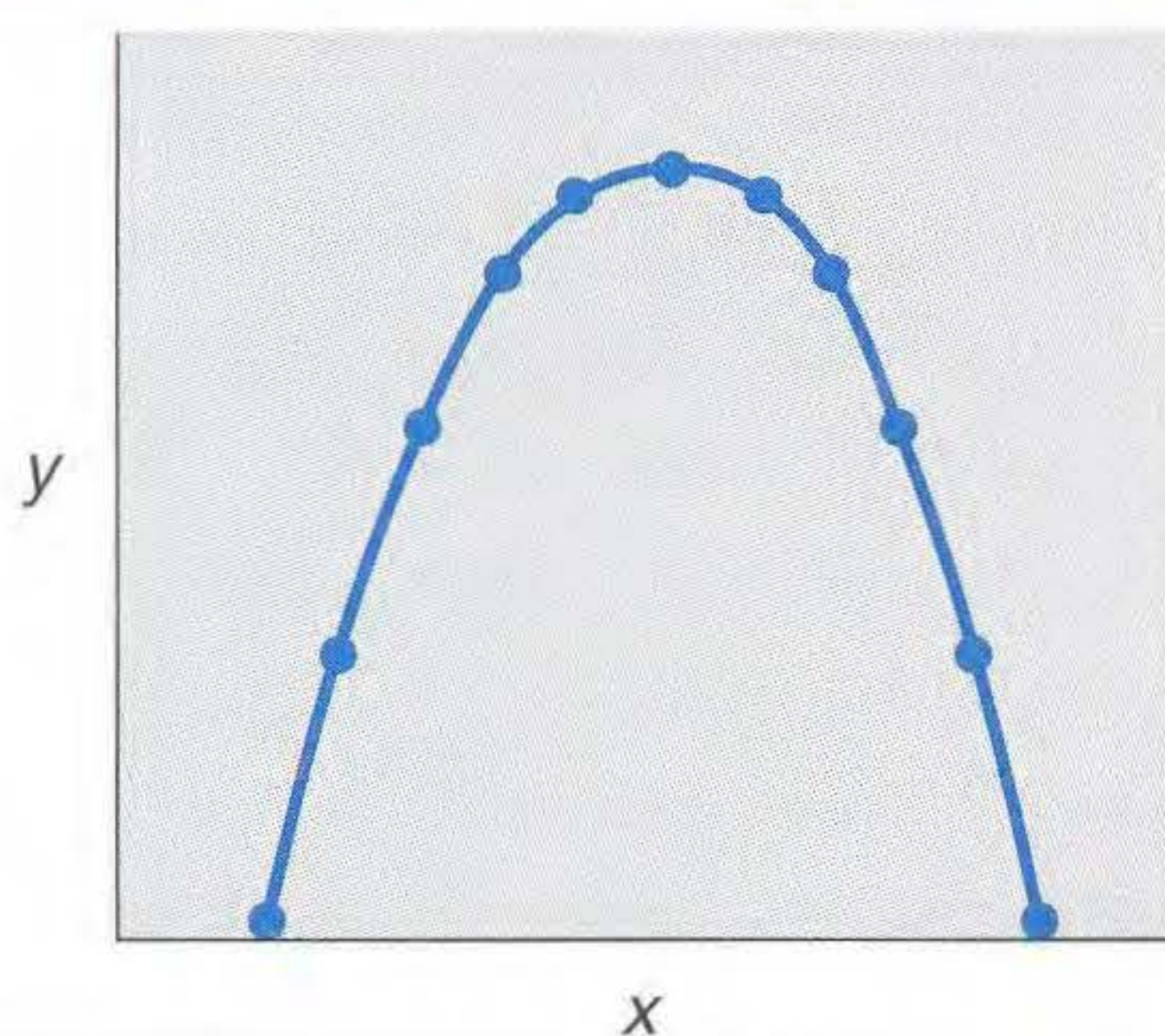


FIG. 18.22 Cada valor de y é perfeitamente previsível pelo valor de x , mas não há correlação linear.

mas, em média, podem ser bem diferentes dos pais como resultado de uma mudança em sua situação social.

Covariança e a variância de uma soma No Quadro 18.3, as variações dos pentes sexuais nas pernas direita e esquerda são 0,8275 e 1,1475, que somam 1,975, mas a variância da soma das duas pernas T é de apenas 1,71. Isto é, a variância do total é menor que a soma das variâncias das partes. Essa discrepância é uma consequência da correlação negativa entre os lados esquerdo e direito. Lados esquerdos maiores estão associados a lados direitos menores e vice-versa, e, assim, a soma dos dois lados varia menos que a soma das variâncias dos dois lados separados. Se, por outro lado, houver uma correlação positiva entre os lados, então os lados esquerdo e direito maiores ocorreriam juntos e a variação da soma dos dois lados seria maior que a soma das duas variâncias separadas. Em geral, se $x + y = T$, então

$$s_T^2 = s_x^2 + s_y^2 + 2 \text{ cov } xy$$

Para os dados do Quadro 18.3,

$$s_T^2 = 1,71 = 0,8275 + 1,1475 - 2(0,1325)$$

Regressão O coeficiente de regressão nos fornece apenas uma estimativa da *precisão* da relação entre duas variáveis. Um problema relacionado é prever o valor de uma variável considerando o valor da outra. Se x aumenta em duas unidades, quanto y aumentará? Se duas variáveis estão linearmente correlacionadas, então a relação pode ser expressa como

$$y = bx + a$$

onde b é a inclinação da linha relacionando y a x e a é o intercepto y dessa linha.

A Figura 18.23 mostra um diagrama de pontos para duas variáveis, y e x , juntos com uma linha reta expressando a

Uma linha de regressão expressa a tendência linear geral de y com o aumento de x

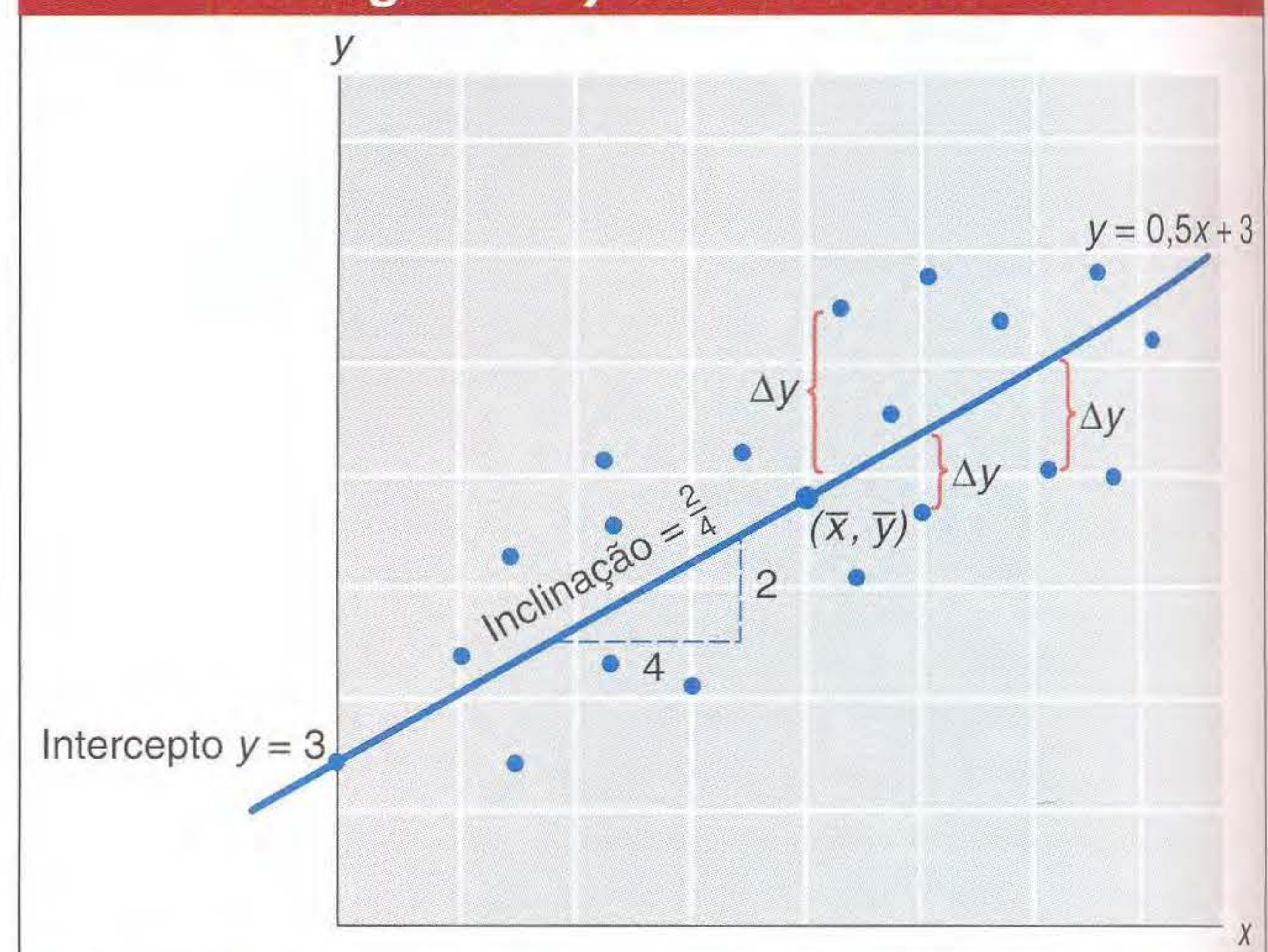


FIG. 18.23 Um diagrama disperso mostrando a relação entre duas variáveis, x e y , com a linha de regressão de y em x . Essa linha, com uma inclinação de $2/4$, minimiza os quadrados dos desvios (Δy).

tendência linear de y com o aumento de x . Essa linha, chamada de **linha de regressão de y em x** , foi posicionada de modo que os desvios dos pontos da linha são o menor possível. Especificamente, se Δy é a distância de qualquer ponto da linha na direção y , então a linha foi escolhida de modo que a soma de $(\Delta y)^2$ é igual a um mínimo. Qualquer outra linha reta passada pelos pontos no diagrama disperso terá um desvio quadrado total dos pontos dela.

Nós não podemos encontrar essa **linha de regressão dos menores quadrados** por ensaio e erro. Ocorre, entretanto, que, se a inclinação b da linha é calculada por

$$b = \frac{\text{cov } xy}{s_x^2},$$

e se a é então calculado como

$$a = \bar{y} - b\bar{x}$$

de modo que a linha passe pelo ponto \bar{x}, \bar{y} , então esses valores de b e a produzirão uma equação linear da linha de regressão dos menores quadrados.

Note que a equação de previsão precedente não pode prever y exatamente para determinado x , porque há uma dispersão ao redor da linha de regressão dos menores quadrados. A equação prevê o y médio para determinado x se for obtida uma amostra grande o suficiente.

Amostras e populações As seções anteriores descreveram as distribuições de, e algumas estatísticas para, conjuntos particulares de indivíduos que foram colhidas de alguns experimentos ou conjuntos de observações. Para muitas finalidades, entretanto, não estamos realmente interessados nos 100 alunos de graduação ou 18 cobras que foram medidas. Em vez disso, estamos interessados em um mundo mais amplo de fenômenos dos quais esses indivíduos particulares são representativos. Por exemplo, podemos querer saber a

altura média, em geral, dos alunos de graduação dos EUA. Isto é, estamos interessados nas características de um **universo**, do qual nossa pequena coleção de observações é apenas uma **amostra**. As características de uma amostra particular qualquer não são idênticas às do universo, mas variam de amostra para amostra.

Podemos usar a média da amostra para estimar a verdadeira média do universo, mas a variância da amostra e a covariância estarão na média um pouco menor que o valor real no universo. Isso é porque os desvios da média da amostra não são todos independentes uns dos outros. De fato, por definição, a soma de todos os desvios das observações da média é 0. (Tente provar como exercício.) Portanto, se soubermos de $N - 1$ dos desvios da média em uma amostra de N observações, podemos calcular o desvio que falta, porque todos os desvios devem somar zero.

Podemos facilmente corrigir essa tendenciosidade em nossa estimativa de variância. Sempre que estivermos interessados na variância de um conjunto de medidas, não como uma característica da amostra particular, mas como uma estimativa de um universo que a amostra representa, então a quantidade apropriada a usar, em vez de s^2 em si, é $[N/(N - 1)] s^2$. Note que essa nova quantidade é equivalente a dividir a soma dos desvios quadrados por $N - 1$ em vez de N ; logo,

$$\begin{aligned} \left(\frac{N}{N - 1} \right) s^2 &= \left(\frac{N}{N - 1} \right) \frac{1}{N} \sum (x_i - \bar{x})^2 \\ &= \frac{1}{N - 1} \sum (x_i - \bar{x})^2 \end{aligned}$$

Todas essas considerações sobre tendenciosidade aplicam-se à covariância da amostra. Na fórmula precedente para o coeficiente de correlação, entretanto, o fator $N/(N - 1)$ apareceria tanto no numerador quanto no denominador e, portanto, seria cancelado, e podemos ignorá-lo para fins de computação.

► Resumo

Muitas — talvez a maioria — das características fenotípicas que observamos nos organismos variam continuamente. Em muitos casos, a variação da característica é determinada por mais de um único locus segregante. Cada um desses loci pode contribuir igualmente para um fenótipo particular, mas é mais provável que contribuam igualmente. A medida desses fenótipos e a determinação das contribuições de alelos específicos para a distribuição devem ser feitas com base estatística nesses casos. Algumas dessas variações de fenótipo (tais como altura em algumas plantas) podem apresentar uma distribuição normal ao redor de um valor médio; outras (tais como peso das sementes em algumas plantas) ilustrarão uma distribuição desviada ao redor de um valor médio.

Uma característica quantitativa é aquela para a qual as diferenças fenotípicas médias entre os genótipos são pequenas comparadas com a variação entre os indivíduos dentro dos genótipos. Essa situação pode ser verdadeira mesmo para características que são influenciadas por alelos em um

locus. A distribuição de ambientes é refletida biologicamente como uma distribuição de fenótipos. A transformação da distribuição ambiental em distribuição fenotípica é determinada pela norma de reação. As normas de reação podem ser caracterizadas em organismos nos quais grandes números de indivíduos geneticamente idênticos podem ser produzidos. As características são familiares se são comuns aos membros da mesma família, por qualquer motivo. As características são herdáveis, mas apenas se a similaridade surgir de genótipos comuns. Em organismos experimentais, as similaridades ambientais podem ser prontamente distintas das similaridades genéticas, ou herdabilidade. Em humanos, entretanto, é muito difícil determinar se uma característica particular é herdável. Os estudos de normas de reação mostram apenas pequenas diferenças entre genótipos, e essas diferenças não são consistentes em uma ampla gama de ambientes. Assim, genótipos “superiores” em animais domesticados e plantas cultivadas só podem ser superiores em alguns ambientes. Se

ocorrer que os humanos exibem variação genética para várias características mentais e emocionais, essa variação provavelmente favorece um genótipo em relação a outro em uma gama de ambientes.

A tentativa de quantificar a influência dos genes em determinada característica levou à determinação da herdabilidade no sentido amplo (H^2). Em geral, a herdabilidade de uma característica é diferente em cada população e em cada conjunto de ambientes, e, assim, a herdabilidade não pode ser extrapolada de uma população e conjunto de ambientes para outro. Como H^2 caracteriza populações presentes apenas em ambientes presentes, é fundamentalmente errada como

instrumento preditivo. A herdabilidade no sentido restrito (h^2) mede a proporção de variação fenotípica que resulta de substituir um alelo por outro. Essa quantidade, se grande, prevê que a seleção para uma característica ocorrerá rapidamente. Se h^2 é pequena, são necessárias formas especiais de seleção.

Com o uso de cromossomos geneticamente marcados, é possível determinar as contribuições relativas dos diferentes cromossomos para a variação de uma característica quantitativa, para observar a dominância e a epistasia de cromossomos inteiros, e, em alguns casos, mapear genes que estão segregando quanto a uma característica.

► Termos-chave

amostra	função de distribuição	método híbrido-endogâmico
análise de variância	gene candidato	moda
coeficiente de correlação	gene marcador	norma de reação
correlação	genética quantitativa	relação
correlação fenotípica	herdabilidade	resposta de seleção
correlação genética	herdabilidade no sentido amplo (H^2)	seleção familiar
covariância	herdabilidade no sentido restrito (h^2)	seleção truncada
desvio padrão	hipótese de fator múltiplo	tendência central
diferencial de seleção	histograma de frequência	universo
dispersão	linha de regressão de y em x	valor médio parental
distribuição bimodal	linha de regressão dos menores	variação genética aditiva
distribuição estatística	quadrados	variança
efeito aditivo	linhagem segregante	variança ambiental
endogamia	locus de característica quantitativa	variança de dominância
familiaridade	(QTL)	variança genética
frequência relativa	média	

► Problemas Resolvidos

Problema resolvido 1. Em algumas espécies de pássaros canoros, as populações que vivem em regiões geográficas diferentes cantam “dialetos locais” diferentes. Algumas pessoas acreditam que essa diferença em dialetos resulta de diferenças genéticas entre populações, enquanto outros acreditam que essas diferenças surgem de idiosincrasias puramente individuais nos fundadores dessas populações e foram transmitidas de geração em geração por aprendizagem. Crie um programa experimental que determine a importância de fatores genéticos e não-genéticos e suas interações nessa variação de dialetos. Se há uma evidência de diferença genética, que experimentos podem ser feitos para dar uma descrição detalhada do sistema genético, incluindo o número de genes segregantes, suas relações de ligação e seus efeitos fenotípicos aditivos e não-aditivos?

SOLUÇÃO

Esse exemplo foi escolhido porque ele ilustra as consideráveis dificuldades experimentais que surgem quando tentamos examinar afirmações de que as diferenças observadas em

características quantitativas em algumas espécies têm uma base genética. Para ser capaz de dizer algo sobre os papéis dos genes e ambientes desenvolvimentais, é necessário, no mínimo, que os organismos possam ser criados a partir dos zigotos em ambiente laboratorial controlado. Para ser capaz de fazer declarações mais detalhadas sobre os genótipos subjacentes à variação na característica, é necessário ainda que os resultados de cruzamentos entre genitores de fenótipos conhecidos e ancestrais conhecidos sejam observáveis, e que a prole de alguns desses cruzamentos seja, por sua vez, cruzada com outros indivíduos de fenótipo e ancestralidade conhecidos. Muito poucas espécies animais satisfazem a esse requisito, embora seja muito mais fácil fazer cruzamentos controlados em plantas. Vamos supor que a espécie canora em questão possa ser criada e cruzada em cativeiro, mas essa é uma grande suposição.

a. Para determinar se há qualquer diferença genética subjacente à diferença fenotípica observada no dialeto entre as populações, precisamos criar as aves de cada população, desde o ovo, na ausência de estímulos auditivos de seus próprios ancestrais e em várias combinações de ambientes auditivos

de outras populações. Podemos fazer isso criando as aves desde o ovo que foram agrupadas do seguinte modo:

- (1) Em isolamento
- (2) Rodeada de eclosões que consistem apenas em aves derivadas da mesma população
- (3) Rodeada de eclosões que consistem em aves derivadas de outras populações
- (4) Na presença de adultos cantores de uma das outras populações
- (5) Na presença de adultos cantores de sua própria população (como um controle das condições de criação)

Se não houver diferenças genótípicas e todas as diferenças de dialetos forem aprendidas, então as aves no grupo 5 cantarão o dialeto de sua população e aquelas no grupo 4 cantarão o dialeto exógeno. Os grupos 1, 2 e 3 podem não cantar; podem cantar uma canção generalizada não correspondendo a nenhum dos dialetos; ou podem todas cantar o mesmo dialeto. Esse dialeto representaria então o programa de desenvolvimento “intrínseco” não modificado pela aprendizagem.

Se as diferenças de dialetos forem totalmente determinadas por diferenças genéticas, as aves nos grupos 4 e 5 cantarão o mesmo dialeto, o de seus genitores. As aves nos grupos 1, 2 e 3, se cantarem, cada uma cantará o dialeto de sua população parental, independentemente de outras aves em seu grupo. Existem então as possibilidades de resultados menos claros, indicando que ambas as diferenças, genéticas e aprendidas, influenciam a característica. Por exemplo, as aves no grupo 4 podem cantar uma canção com elementos de ambas as populações. Note que, se as aves no grupo controle (grupo 5) não cantam seu dialeto normal, o resto dos resultados não pode ser interpretado, pois as condições artificiais de criação estão interferindo no programa normal de desenvolvimento.

b. Se os resultados dos primeiros experimentos mostram alguma herdabilidade no sentido amplo, então é possível uma outra análise. Essa análise requer uma população geneticamente segregante, feita de um cruzamento entre o dialeto de duas populações, digamos, A e B. Um cruzamento entre os machos da população A e as fêmeas da população B e o cruzamento recíproco darão uma estimativa do grau médio de dominância dos genes que influenciam a característica e indicarão se há alguma ligação ao sexo. (Lembre que, nas aves, a fêmea é o sexo heterogamético.) A prole desse cruzamento e a de todos os cruzamentos subsequentes *devem* ser criadas em condições que não confundam os componentes aprendidos e genéticos das diferenças, como revelado nos experimentos na parte a. Se os efeitos aprendidos não podem ser separados, essa posterior análise genética é impossível.

c. Para localizar os genes que influenciam as diferenças de dialetos, seria necessário um grande número de marcadores genéticos segregantes. Esses marcadores poderiam ser mutantes morfológicos ou variantes moleculares, tais como polimorfismos de sítio de restrição. As famílias que segregam diferenças de características quantitativas seriam examinadas para ver se há co-segregação de algum dos loci marcadores com a característica quantitativa. Esses loci co-segregados

seriam então candidatos a loci ligados aos loci de características quantitativas. Outros cruzamentos entre indivíduos com e sem marcadores mutantes e medidas dos valores de características quantitativas nos indivíduos de F_2 estabeleceriam se há uma ligação real entre marcador e os loci da característica quantitativa. Na prática, é muito improvável que tais experimentos possam ser feitos, em uma espécie de aves canoras, devido ao imenso tempo e esforço necessários para estabelecer linhagens portadoras do grande número de genes marcadores diferentes e polimorfismos moleculares.

Problema resolvido 2. Duas linhagens endocruzadas de feijões são inter cruzadas. Na F_1 , a variância no peso do feijão é medida como 1,5. A F_1 é autofecundada; na F_2 , a variância no peso do feijão é 6,1. Avalie a herdabilidade ampla do peso do feijão na população F_2 desse experimento.

SOLUÇÃO

O importante aqui é reconhecer que toda a variância na população F_1 deve ser ambiental, pois todos os indivíduos devem ter genótipo idêntico. Além disso, a variância de F_2 deve ser uma combinação de componentes ambientais e genéticos, pois todos os genes que são heterozigotos na F_1 irão segregar na F_2 para dar uma disposição de genótipos diferentes que se relacionem com o peso do feijão. Assim, podemos estimar

$$s_e^2 = 1,5$$

$$s_e^2 + s_g^2 = 6,1$$

Portanto,

$$s_g^2 = 6,1 - 1,5 = 4,6$$

e a herdabilidade ampla é

$$H^2 = \frac{4,6}{6,1} = 0,75 \text{ (75\%)}$$

Problema resolvido 3. Em uma população experimental de *Tribolium* (besouro da farinha), o tamanho do corpo mostra uma distribuição contínua com uma média de 6 mm. Um grupo de machos e fêmeas com tamanho de corpo de 9 mm são removidos e cruzados. Os tamanhos de corpo de sua prole têm uma média de 7,2 mm. A partir desses dados, calcule a herdabilidade no sentido restrito do tamanho do corpo nessa população.

SOLUÇÃO

A seleção diferencial é $9 - 6 = 3$ mm, e a resposta de seleção é $7,2 - 6 = 1,2$ mm. Portanto, a herdabilidade no sentido restrito é

$$h^2 = \frac{1,2}{3} = 0,4 \text{ (40\%)}$$

► Problemas

PROBLEMAS BÁSICOS

- 1. Distinga a variação contínua e descontínua em uma população, e cite alguns exemplos de cada.
- 2. O seguinte quadro mostra a distribuição do número de cerdas em *Drosophila*. Calcule a média, a variância e o desvio padrão dessa distribuição.

Nº de cerdas	Nº de indivíduos
1	1
2	4
3	7
4	31
5	56
6	17
7	4

- 3. Um livro sobre o problema da herdabilidade de QI faz as seguintes três afirmativas. Discuta a validade de cada afirmativa e suas implicações sobre a compreensão do autor de h^2 e H^2 .
 - a. “A pergunta interessante então é ... ‘O quanto é herdável?’ A resposta [0,01] tem uma aplicação teórica e prática muito diferente da resposta [0,99].” (Os autores estão falando de H^2 .)
 - b. “Empiricamente, quando a questão é educacional, H^2 é o coeficiente mais relevante, e, quando estão sendo discutidas a eugenia e a disgenia (reprodução de indivíduos selecionados), h^2 geralmente é mencionado.”
 - c. “Mas, se diferentes padrões de habilidade são derivados de diferenças em genes ... não é relevante avaliar discriminações em contratação. O que pode ser relevante é decidir o que, em longo prazo, pode ser feito para mudar a situação.”(O problema 3 é de J.C. Loehlin, G. Lindzey, and J.N. Spuhler, *Race Differences in Intelligence*. Copyright 1975 by W.H. Freeman and Company.)
- 4. Usando os conceitos de normas de reação, distribuição ambiental, distribuição genotípica e distribuição fenotípica, tente refazer a seguinte frase em termos mais exatos: “Oitenta por cento da diferença no desempenho de QI entre os dois grupos é genética”. O que significa falar de herdabilidade de uma diferença entre dois grupos?

PROBLEMAS DESAFIADORES

- 5. Em um grande rebanho de gado, três características diferentes mostrando distribuição contínua são medidas, e são calculadas as variâncias no seguinte quadro:

Variância	Características		
	Tamanho da perna	Tamanho do pescoço	Conteúdo de gordura
Fenotípico	310,2	730,4	106,0
Ambiental	248,1	292,2	53,0
Adicional genético	46,5	73,0	42,4
Dominância genética	15,6	365,2	10,6

- a. Calcule as herdabilidades no sentido amplo e no sentido restrito para cada característica.
- b. Na população de animais estudada, que característica responderia melhor à seleção? Por quê?
- c. É feito um projeto para diminuir a média de conteúdo de gordura no rebanho. A média do conteúdo de gordura é atualmente de 10,5%. Os animais com 6,5% de gordura são entrecruzados como genitores da geração seguinte. Que média de conteúdo de gordura pode ser esperada nos descendentes desses animais?
- 6. Suponha que dois heterozigotos triplos $A/a ; B/b ; C/c$ sejam cruzados. Suponha que os três loci estejam em cromossomos diferentes.
 - a. Que proporção da prole são homozigotos em um, dois e três loci, respectivamente?
 - b. Que proporção da prole leva 0, 1, 2, 3, 4, 5 e 6 alelos (representados por letras maiúsculas), respectivamente?
- 7. No problema 6, suponha que o efeito fenotípico médio dos três genótipos no locus A seja $A/A = 4$, $A/a = 3$ e $a/a = 1$ e que existam efeitos similares para os loci B e C . Além disso, suponha que os efeitos dos loci sejam somados. Calcule e faça um gráfico da distribuição de fenótipos na população (supondo ausência de variância ambiental).
- 8. No problema 7, suponha que haja um limiar na característica fenotípica de modo que, quando o valor fenotípico está acima de 9, uma *Drosophila* individual tem três cerdas; quando o valor fenotípico está entre 5 e 9, o indivíduo tem duas cerdas; e, quando o valor é 4 ou menos, a mosca tem uma cerda. Descreva o resultado dos cruzamentos dentro e dentre as classes de cerdas. Considerando o resultado, você pode deduzir a situação genética subjacente?
- 9. Suponha que a forma geral de distribuição de uma característica para determinado genótipo é

$$f = 1 - \frac{(x - \bar{x})^2}{s_e^2}$$

na faixa de x onde f é positivo.

a. Na mesma escala, plote as distribuições para três genótipos com as seguintes médias e variâncias ambientais.

Genótipo	\bar{x}	s_e^2	Faixa aproximada de fenótipo
1	0,20	0,3	$x = 0,03$ a $x = 0,37$
2	0,22	0,1	$x = 0,12$ a $x = 0,24$
3	0,24	0,2	$x = 0,10$ a $x = 0,38$

b. Plote a distribuição fenotípica que resultaria se os três genótipos fossem igualmente freqüentes em uma população. Você pode imaginar modos diferentes? Caso sim, quais são eles?

10. Os seguintes conjuntos de dados hipotéticos representam observações de duas variáveis (x e y). Plote cada conjunto de pares de dados como um diagrama disperso. Procure na plotagem dos pontos, e faça uma suposição intuitiva sobre a correlação entre x e y . Calcule então o coeficiente de correlação para cada conjunto de pares de dados, e compare esse valor com sua estimativa.

- a. (1, 1); (2, 2); (3, 3); (4, 4); (5, 5); (6, 6).
- b. (1, 2); (2, 1); (3, 4); (4, 3); (5, 6); (6, 5).
- c. (1, 3); (2, 1); (3, 2); (4, 6); (5, 4); (6, 5).
- d. (1, 5); (2, 3); (3, 1); (4, 6); (5, 4); (6, 2).

11. Descreva um protocolo experimental para estudos de parentes que possa estimar a herdabilidade ampla do alcoolismo. Lembre que você deve fazer uma definição observacional da própria característica.

12. Uma linhagem selecionada pelo alto número de cerdas em *Drosophila* tem uma média de 25 cerdas esternopleurais, enquanto uma linhagem selecionada para baixo tem uma média apenas de 2. Os estoques marcadores para os dois grandes autossomos II e III são usados para criar estoques com várias misturas de cromossomos das linhagens alto (h) e baixo (l). O número médio de cerdas para cada combinação cromossômica é o seguinte:

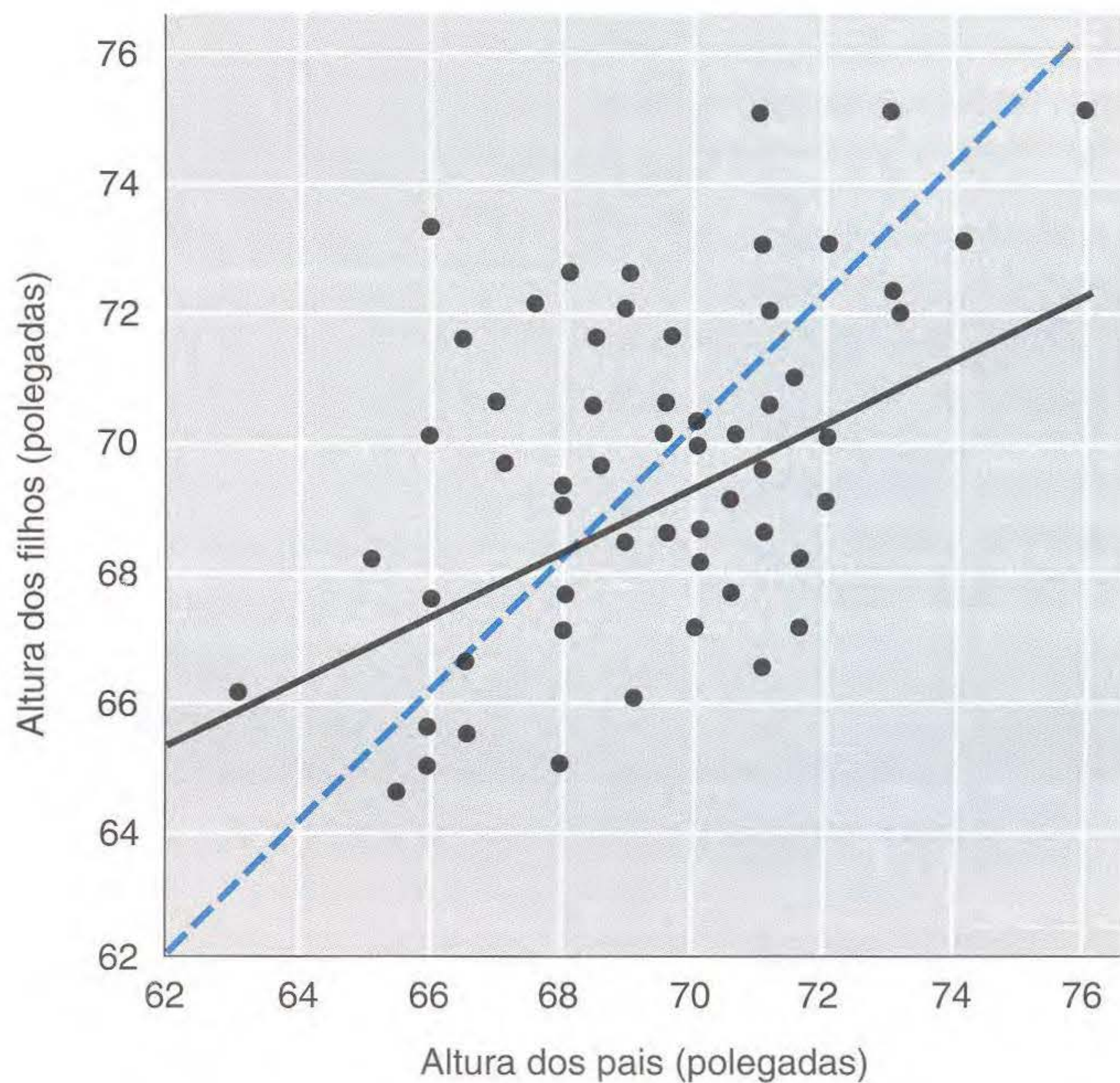
$\frac{h}{h} \frac{h}{h}$ 25,1	$\frac{h}{l} \frac{h}{h}$ 22,2	$\frac{l}{l} \frac{h}{h}$ 19,0
$\frac{h}{h} \frac{h}{l}$ 23,0	$\frac{h}{l} \frac{h}{l}$ 19,9	$\frac{l}{l} \frac{h}{l}$ 14,7
$\frac{h}{h} \frac{l}{l}$ 11,8	$\frac{h}{l} \frac{l}{l}$ 9,1	$\frac{l}{l} \frac{l}{l}$ 2,3

13. Suponha que o número de facetas é medido em uma população de *Drosophila* sob várias condições de temperatura. Suponha ainda que é possível estimar a variância genética total (s_g^2) bem como a distribuição fenotípica.

Finalmente, suponha que existem apenas dois genótipos na população. Desenhe os pares de normas de reação que levariam aos seguintes resultados:

- a. Um aumento na temperatura média diminui a variância fenotípica.
- b. Um aumento na temperatura média aumenta H^2 .
- c. Um aumento na temperatura média aumenta (s_g^2), mas diminui H^2 .
- d. Um aumento na *variância* de temperatura muda uma distribuição fenotípica unimodal em bimodal (uma norma de reação é suficiente aqui).

14. Francis Galton comparou as alturas de estudantes de graduação com as alturas de seus pais, com os resultados no gráfico seguinte.



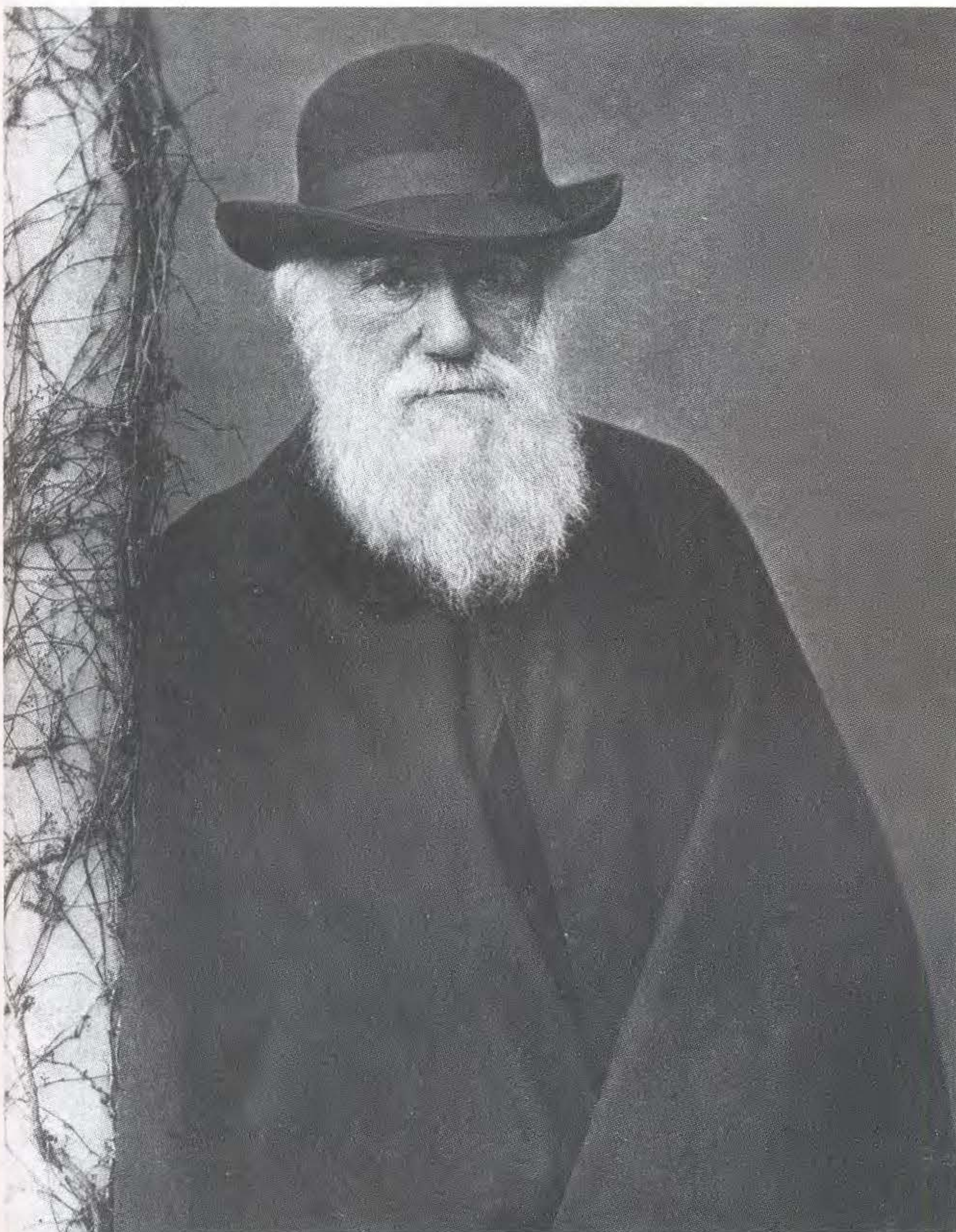
A altura média de todos os pais é a mesma que a altura média de todos os filhos, mas as classes individuais de altura não são iguais ao longo das gerações. Os pais bem altos tinham filhos um pouco mais baixos, enquanto os pais mais baixos tinham filhos um pouco mais altos. Como resultado, a melhor linha que se pode obter através dos pontos no diagrama de dispersão tem uma inclinação de cerca de 0,67 (linha contínua) em vez de 1,00 (linha tracejada). Galton usou o termo *regressão* para descrever essa tendência de o fenótipo dos filhos ser mais próximo da média da população que o fenótipo da média da população.

- a. Proponha uma explicação para esta regressão.
- b. Como a regressão e a herdabilidade estão relacionadas aqui?

(Gráfico de W.F. Bodmer and L.L. Cavalli-Sforza, *Genetics, Evolution, and Man*. Copyright 1976 de W.H. Freeman and Company.)

Genética Evolutiva

19



Charles Darwin. [CORBIS/Bettmann.]

A teoria moderna da evolução é tão identificada com o nome de Charles Darwin (1809-1882) que muitas pessoas pensam que foi Darwin quem primeiro formulou o conceito de que os organismos evoluíam, mas isso não é verdade. Muitos acadêmicos abandonaram a noção de espécies fixas, inalteradas desde sua origem em uma grande criação da vida, muito antes da publicação do *The Origin of Species* de Darwin, em 1859. Nessa época, a maioria dos biólogos concordavam que novas espécies surgiam por algum processo de evolução de espécies anteriores. O problema era explicar *como* essa evolução ocorreria.

Darwin deu uma explicação detalhada do mecanismo do processo evolutivo. A teoria de Darwin do mecanismo da evolução começou com a variação que existe entre os organismos dentro de uma espécie. Os indivíduos de uma geração são qualitativamente diferentes uns dos outros. A evolução da espécie como um todo resulta do fato de que vários tipos diferem em suas taxas de sobrevivência e reprodução, e, assim, as frequências relativas dos tipos mudam

Perguntas Fundamentais

- Quais são os princípios básicos do mecanismo darwiniano da evolução?
- Quais são os papéis da seleção natural e outros processos na evolução, e como eles interagem uns com os outros?
- Como as características evoluem?
- Quais são as assinaturas da seleção natural ou sua ausência nas seqüências de DNA?
- Por que as seqüências reguladoras são importantes na evolução de características morfológicas?

Tópicos

- 19.1 Evolução darwiniana
- 19.2 Uma síntese de forças: variação e divergência de populações
- 19.3 Vários picos adaptativos
- 19.4 Variação genética
- 19.5 Mutação e evolução molecular
- 19.6 Relação entre genética e mudança funcional na evolução de proteínas
- 19.7 Evolução regulatória
- 19.8 A origem de novos genes
- 19.9 Evidência genética de ancestralidade comum na evolução
- 19.10 O processo de especiação

com o tempo. A evolução, desse ponto de vista, é um processo de variação.

Para Darwin, a evolução do grupo resultou da sobrevivência diferencial e reprodução de variantes individuais *já existentes* no grupo — variantes que surgem de um modo não-relacionado ao ambiente mas cuja sobrevivência e reprodução dependem do ambiente.

Darwin criou uma nova explicação para o fenômeno aceito da evolução. Ele disse que a população de determinada espécie em determinada época incluía indivíduos com características variadas. A população da geração seguinte continha uma frequência maior desses tipos que, bem-sucedidamente, sobrevivem e se reproduzem nas condições ambientais existentes. Assim, as frequências de vários tipos dentro das espécies mudariam com o tempo.

Há uma similaridade óbvia entre o processo de evolução como Darwin o descreveu e o processo pelo qual o criador de animais e plantas melhora um estoque doméstico. Um criador de plantas seleciona as de maior produtividade a partir da população atual e (tanto quanto possível) as usa como genitoras da geração seguinte. Se as características que causam a maior produção são herdáveis, então a geração seguinte deve ter uma grande produtividade. Não foi por acidente que Darwin escolheu o termo **seleção natural** para descrever seu modelo de evolução pelas taxas diferenciais de reprodução de variantes diferentes na população. Como modelo para esse processo evolutivo na natureza, ele teve em mente a seleção que os criadores exerciam em gerações sucessivas de plantas e animais domésticos:

... podemos ter dúvidas (lembrando que nascem muito mais indivíduos do que possivelmente sobrevivem) de que os indivíduos que têm alguma vantagem, mesmo que pequena, sobre os outros terão a melhor chance de sobreviver e de se reproduzir? Por outro lado, podemos ter certeza de que qualquer variação, mesmo que prejudicial em pequeno grau, será rigidamente destruída. A essa preservação de variações favoráveis e rejeição de variações prejudiciais, chamo de Seleção Natural. [No *The Origin of Species*, Capítulo IV].

► 19.1 Evolução Darwiniana

Podemos resumir a teoria da evolução de Darwin pela seleção natural em três princípios:

1. **Princípio da variação.** Entre os indivíduos dentro de qualquer população, há uma variação na morfologia, fisiologia e comportamento.
2. **Princípio da hereditariedade.** A prole assemelha-se a seus genitores mais do que a indivíduos não-aparentados.
3. **Princípio da seleção.** Algumas formas são mais bem-sucedidas em sobreviver e se reproduzir do que outras formas em determinado ambiente.

Claramente, um processo seletivo pode produzir mudança na composição da população apenas se existirem algumas

variações entre as quais selecionar. Se todos os indivíduos forem idênticos, nenhuma quantidade de reprodução diferencial de indivíduos irá alterar a composição da população. Além disso, a variação deve ser em parte herdável se a reprodução diferencial for alterar a composição genética da população. Se animais grandes dentro de uma população tiverem mais prole que animais pequenos mas sua prole em média não for maior do que a de animais pequenos, então não haverá mudança na composição da população de uma geração para outra. Finalmente, se todos os tipos variantes deixam, em média, o mesmo número de prole, então podemos esperar que a população permaneça inalterada.

Os princípios de Darwin da variação, hereditariedade e seleção devem se manter se houver evolução por um mecanismo de variação.

A explicação darwiniana da evolução deve ser capaz de responder por dois aspectos diferentes da história da vida. Um deles é a sucessiva mudança de forma e função que ocorre em uma única linhagem contínua de descendentes: **evolução filética**. A Figura 19.1 mostra tal mudança contínua, em um período de 40 milhões de anos, no tamanho e curvatura da concha esquerda da ostra *Gryphea*. Outro é a **diversificação** que ocorre entre as espécies: na história da vida na Terra, devem ter existido muitas espécies contemporâneas tendo formas bem diferentes e vivendo de modos diferentes. A Figura 19.2 mostra alguma variedade de formas do molusco bivalvo que existiu em várias épocas nos últimos 300 milhões de anos. Cada espécie tornou-se extinta e mais de 99,9% das espécies que existiam já foram extintas, embora o número de espécies e a diversidade de suas formas e funções tenham aumentado nos últimos bilhões de anos. Assim, as espécies não só devem estar mudando, mas devem ter dado origem a novas e diferentes espécies no curso da evolução.

Ambos os processos, evolução filética e diversificação, são consequências de variação herdável dentro das populações. A variação herdável fornece a matéria-prima para sucessivas mudanças dentro de uma espécie e para a multiplicação de novas espécies. Os mecanismos básicos dessas mudanças (como discutido no Capítulo 17) são a origem de nova variação genética por diferentes tipos de processos mutacionais, a mudança na frequência de alelos dentro das populações por processos seletivos e aleatórios, a divergência de populações diferentes devido a forças seletivas diferentes ou devido à deriva aleatória e à redução da variação entre populações por migração. Por esses mecanismos básicos, podem ser derivados conjuntos de princípios que controlam mudanças na composição genética das populações. A aplicação desses princípios de genética de populações fornece uma teoria genética detalhada da evolução.

A evolução, sob o esquema darwiniano, é a conversão da variação herdável entre indivíduos dentro da população em diferenças herdáveis entre populações no tempo e no espaço por mecanismos de genética de populações.

Evolução filética é a mudança sucessiva de forma e função

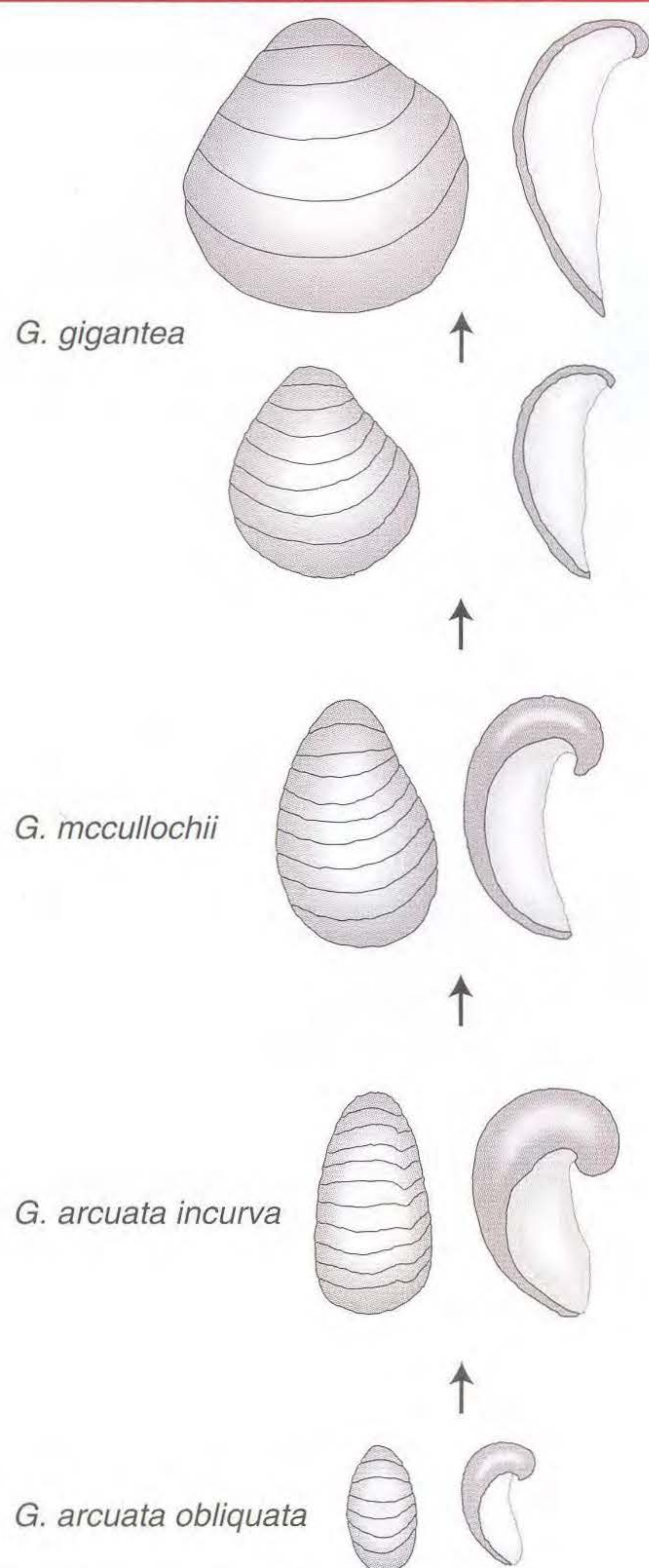


FIG. 19.1 O molusco bivalvo *Gryphaea* sofreu mudanças sucessivas no tamanho e curvatura da concha no curso de sua evolução filética no início do Jurássico. É mostrada apenas a concha esquerda. Em cada caso, a parte traseira da concha e uma seção longitudinal são ilustradas. [De A. Hallam, "Morphology, Palaeoecology and Evolution of the Genus *Gryphaea* in the British Lias," *Philos. Trans. R. Soc. Lond. Ser. B* 254, 1968, 124.]

► 19.2 Uma Síntese de Forças: Variação e Divergência de Populações

Quando Darwin chegou às ilhas Galápagos em 1835, ele encontrou um incrível grupo de pássaros tipo tentilhão que constituíram um caso sugestivo para o desenvolvimento de sua teoria da evolução. O arquipélago de Galápagos é um aglomerado de 29 ilhas e ilhotas de tamanhos diferentes que ficam a cerca de 600 milhas da costa do Equador. Os tentilhões são geralmente comedores de sementes com bicos para quebrar as duras cascas destas. A Figura 19.3 mostra 13 espécies de tentilhão de Galápagos. As espécies de Galápagos, embora claramente sejam tentilhões, exibem uma imensa variação de comportamento alimentar e na forma de bico que correspondem a suas fontes alimentares. Por exemplo, o tentilhão vegetariano de árvores usa seu bico pesado para comer frutos e folhas; o tentilhão insetívoro tem um bico com uma ponta para comer insetos grandes; e, o mais marcante de tudo, o tentilhão pica-pau usa seu bico para obter sua presa de insetos fazendo furos nas árvores. Essa diversidade de espécies surgiu de uma população original de tentilhões comedores de sementes que chegaram às ilhas Galápagos vindos da África do Sul e popularam as ilhas. Os descendentes dos colonizadores originais espalharam-se pelas diferentes ilhas e para partes diferentes das grandes ilhas, formando populações locais que divergiram umas das outras e acaba-

Diversificação é a geração de muitas espécies diferentes

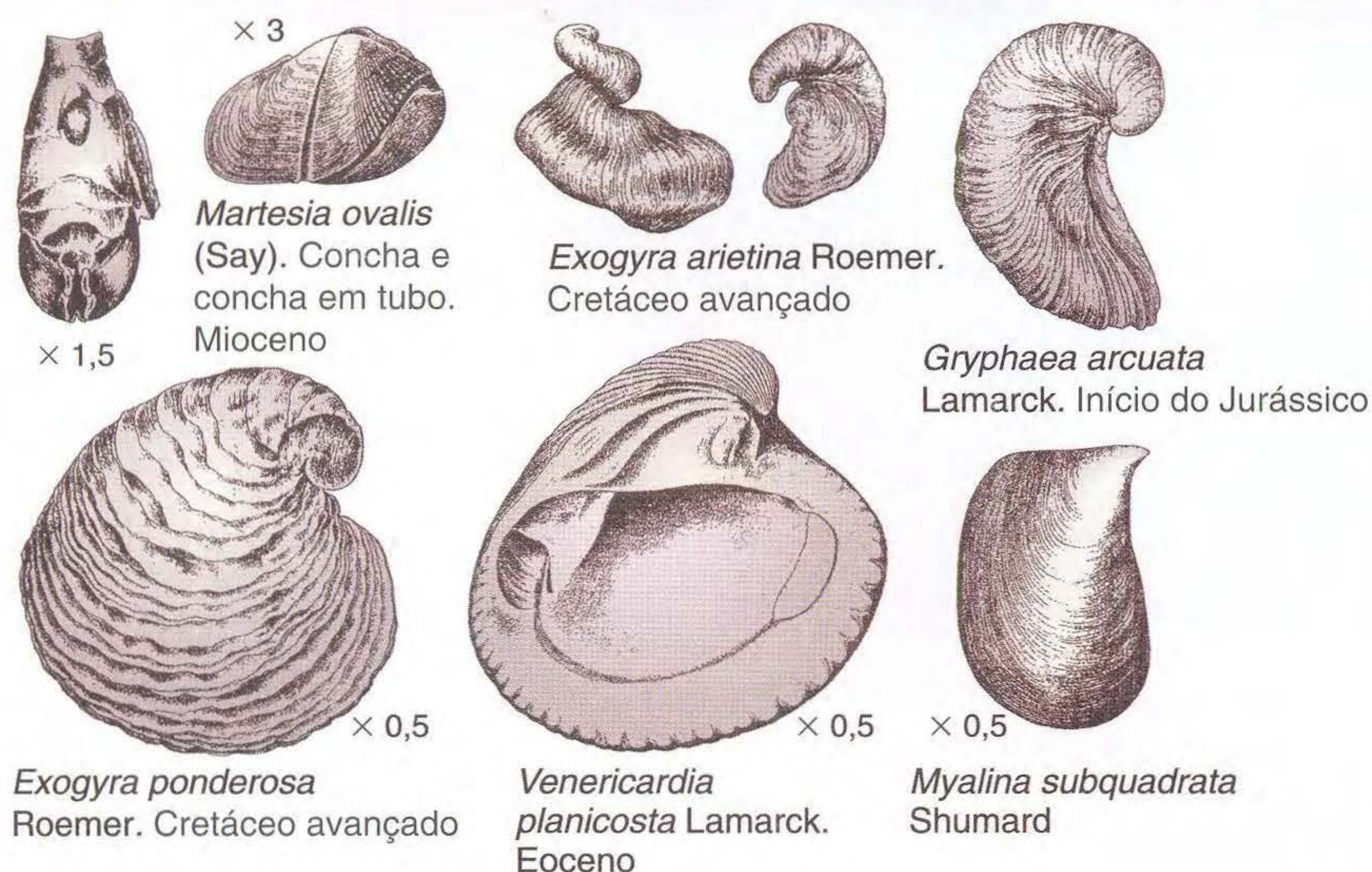


FIG. 19.2 Uma variedade de conchas de moluscos bivalvos apareceu nos últimos 300 milhões de anos de evolução. [De C.L. Fenton and M.A. Fenton, *The Fossil Book*. Doubleday, 1958.]

Uma diversidade de espécies pode resultar de adaptação

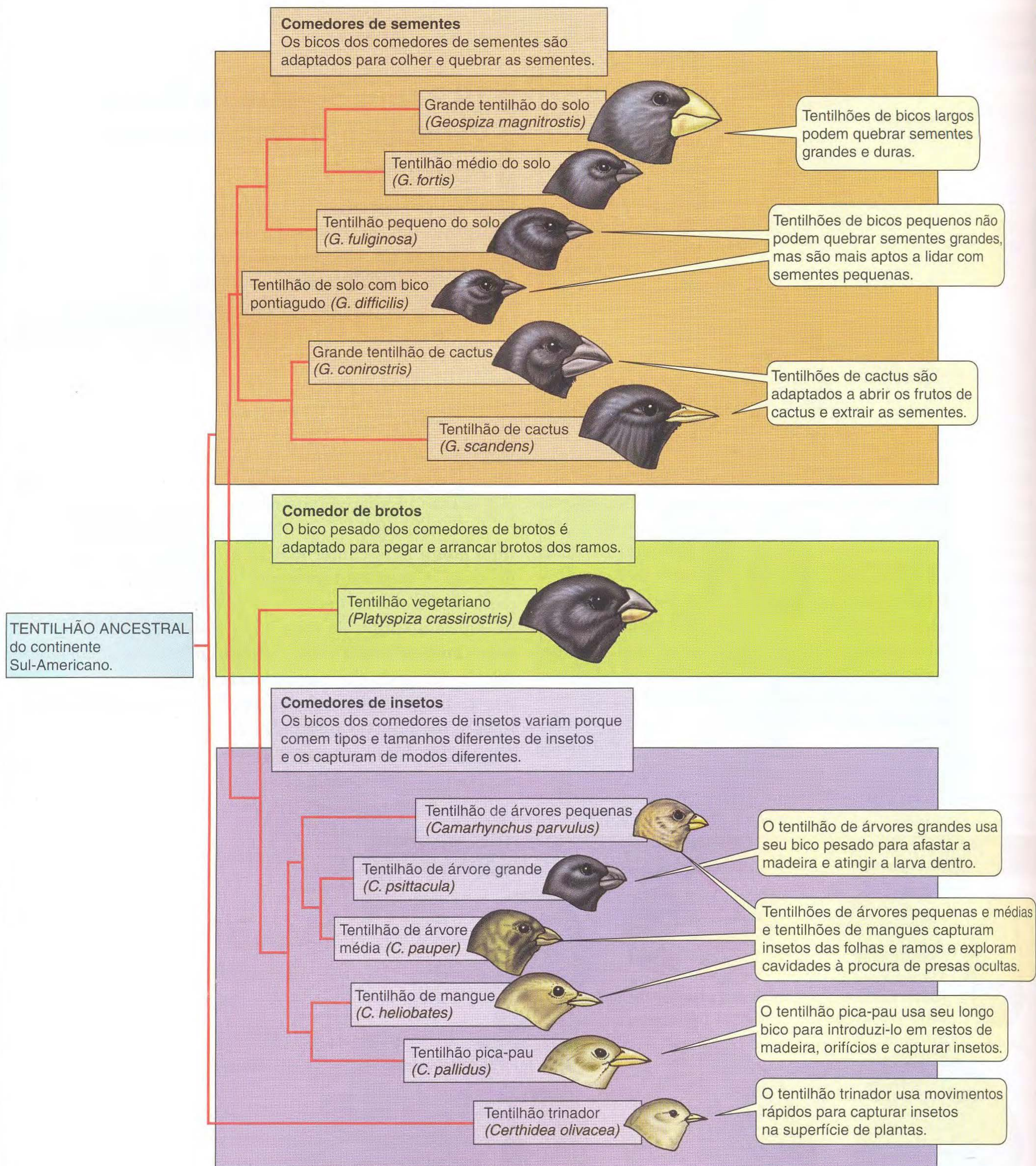


FIG. 19.3 As treze espécies de tentilhões encontrados nas ilhas Galápagos. [De W.K. Purves, G.H. Orians, and H.C. Heller, *Life: The Science of Biology*, 4th ed. Sinauer Associates/W.H. Freeman and Company, 1995, Fig. 20.3, p. 450.]

O intercâmbio de forças evolutivas influencia a variação

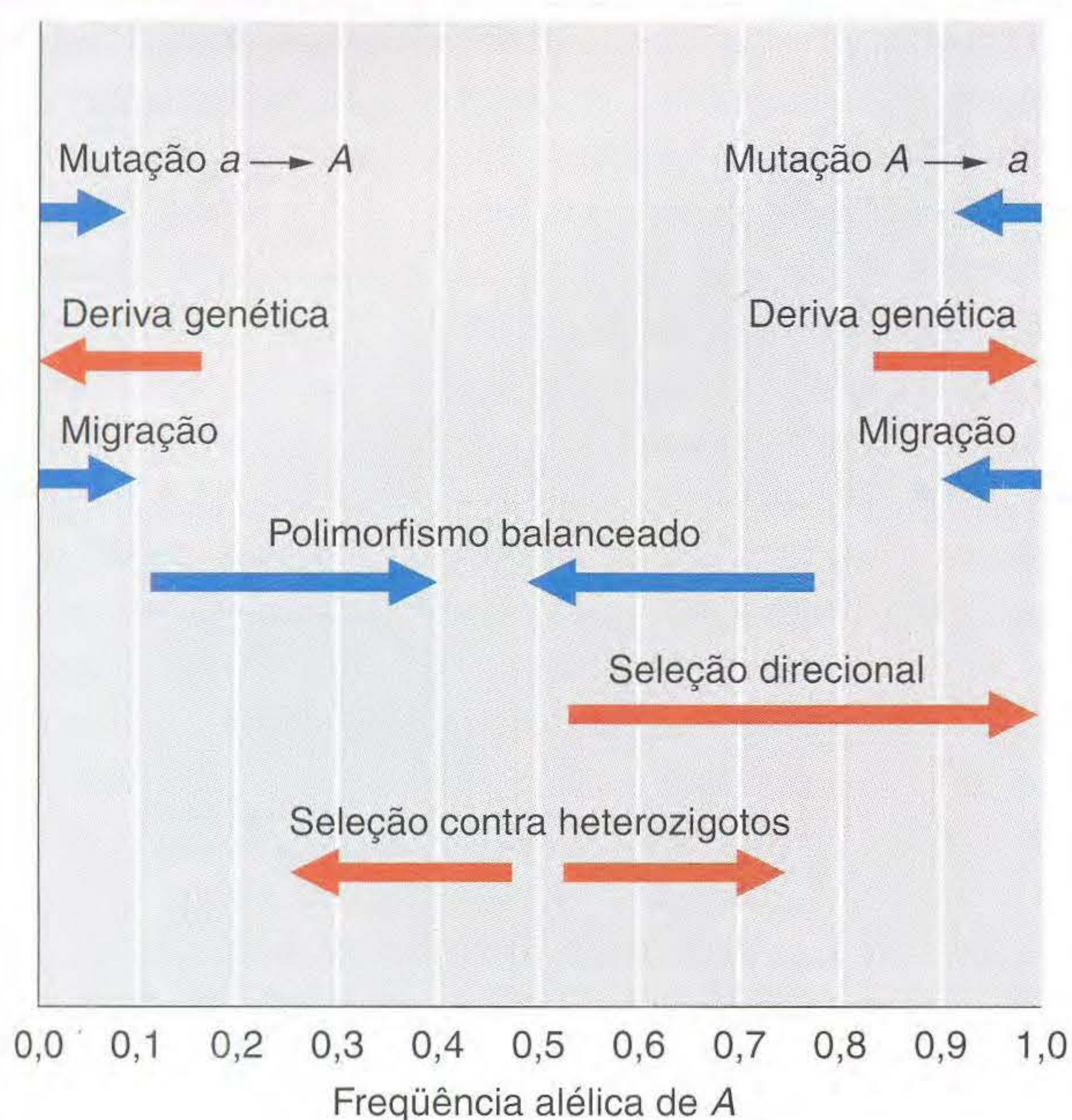


FIG. 19.4 Os efeitos na frequência alélica de várias forças da evolução. As setas azuis mostram uma tendência para aumento de variação dentro da população; as setas vermelhas, uma variação diminuída.

ram formando espécies diferentes. Os tentilhões ilustram dois aspectos da evolução que precisam ser explicados. Como uma espécie original com determinado conjunto de características deu origem a uma diversidade de espécies, cada uma com sua própria forma e função? Como as características das espécies adaptaram-se aos ambientes nos quais vivem as espécies? Esses são os problemas da origem da *diversidade* e a origem da *adaptação*.

A variação genética dentro das populações e entre elas é o resultado do intercâmbio das forças da mutação, migração, seleção e deriva genética discutidos no Capítulo 17 (Figura 19.4). Geralmente, como mostra o Quadro 19.1, as forças que aumentam ou mantêm a variação dentro das populações impedem que as populações venham a divergir umas das outras, enquanto as forças que tornam cada população homozigota fazem com que elas venham a divergir. Assim, a deriva aleatória (ou endogamia) produz homozigose, enquanto faz com que populações diferentes venham a divergir. Essa tendência para a divergência e homozigose é contrabalançada pelo fluxo constante de mutação e migração de indivíduos entre populações, ambas as quais introduzem variação nas populações, tornando-as mais similares.

Considere a situação de um locus geneticamente variável com dois alelos, digamos A e a , nas frequências p e $1 - p$, respectivamente, em uma grande população. Suponha que grupos de populações de ilhas isoladas foram fundados por migrantes dessa única população. Os fundadores originais de cada população são pequenas amostras da população doadora e, assim, diferem uns dos outros em frequências alélicas devido a um efeito de amostragem aleatória. Essa variação inicial é chamada de **efeito do fundador**. Em gerações seguintes, a deriva genética aleatória altera mais as frequên-

Quadro 19.1 Como as Forças da Evolução Aumentam (+) ou Diminuem (-) a Variação dentro de Populações e entre Populações

Força	Variação dentro de populações	Variação entre populações
Endogamia ou deriva genética	-	+
Mutação	+	-
Migração	+	-
Seleção direcional	-	+/-
Balanceamento	+	-
Incompatível	-	+

cias alélicas dentro de cada população. A frequência de cada um dos alelos irá mover-se para 1 ou 0 em cada população, mas a frequência alélica média em toda a população permanecerá constante. Com o passar do tempo, as frequências gênicas entre as populações divergem e algumas se tornam fixadas para um dos alelos. Após $4N$ gerações, 80% das populações estão fixadas, uma proporção p sendo homozigota A/A e uma proporção $1 - p$ sendo homozigota a/a . Eventualmente, todas as populações ficariam fixadas em A/A ou a/a nessas proporções.

O processo de diferenciação por endogamia em populações de ilhas é lento, mas não em uma escala de tempo evolutiva ou geológica. Se uma ilha pode suportar, digamos, 10.000 indivíduos de uma espécie de roedores, então, após 20.000 gerações (cerca de 7.000 anos, supondo 3 gerações por ano), a população será homozigota para cerca de metade de todos os loci que estavam inicialmente no máximo de heterozigose. Além disso, a ilha será diferenciada de outras ilhas similares de dois modos: (1) os loci que são fixados nessa ilha ainda estarão segregando em muitas outras ilhas ou fixados em um alelo diferente e (2) os loci que ainda estão segregando em todas as ilhas irão variar em frequências alélicas de ilha para ilha.

Cada população de cada espécie é finita em tamanho; logo, todas as populações acabarão tornando-se homozigotas e diferenciadas umas das outras como resultado de endogamia. Toda a variação seria eliminada e a evolução terminaria. Na natureza, entretanto, novas variações estão sempre sendo introduzidas nas populações por mutação e por migração entre localidades. Assim, a variação real disponível para a seleção natural é um balanço entre a introdução de nova variação e perda por endogamia local. Lembre-se, do Capítulo 17, que a taxa de perda de heterozigose em uma população fechada é de $1/(2N)$ por geração, e, assim, qualquer diferenciação efetiva entre populações que ocorra devido à deriva será negada se nova variação é introduzida nessa taxa ou uma taxa maior. Se m é a taxa de migração para determinada população e μ é a taxa de mutação para novos alelos por geração, então mais ou menos (para uma ordem de magnitude) uma população reterá a maior parte de sua heterozigose e não se diferenciara muito de outras populações por endogamia local se

$$m \geq \frac{1}{N} \quad \text{ou} \quad \mu \geq \frac{1}{N}$$

ou, em outras palavras, se

$$Nm \geq 1 \text{ ou } N\mu \geq 1$$

Para populações de tamanho intermediário e mesmo de tamanho ligeiramente grande, é improvável que $N\mu \geq 1$. Por exemplo, se o tamanho da população é de 100.000, então, para evitar a perda de variação, a taxa de mutação deve exceder 10^{-5} , o que está um pouco do lado alto para taxas de mutação conhecidas, embora não seja uma taxa desconhecida. Por outro lado, uma taxa de migração de 10^{-5} por geração não é absurdamente grande. De fato

$$m = \frac{\text{número de migrantes}}{\text{tamanho total da população}} = \frac{\text{número de migrantes}}{N}$$

Assim, a exigência de que $Nm \geq 1$ é equivalente à exigência de que

$$Nm = N \times \frac{\text{número de migrantes}}{N} \geq 1$$

ou que

$$\text{o número de indivíduos migrantes} \geq 1$$

independentemente do tamanho da população. Para muitas populações, mais de um único indivíduo migrante por geração é bem provável. As populações humanas (mesmo populações tribais isoladas) têm uma taxa de migração maior do que o valor mínimo, e, como resultado, não é conhecido nenhum locus em humanos para o qual um alelo é fixado em algumas populações e um alelo alternativo é fixado em outras.

Os efeitos da seleção são mais variáveis que os da deriva genética aleatória porque a seleção pode ou não levar uma população para a homozigose. A **seleção direcional** leva uma população para a homozigose, rejeitando a maioria de novas mutações à medida que elas são introduzidas mas ocasionalmente (se a mutação é vantajosa) espalham um novo alelo pela população para criar um novo estado homozigoto. Se a seleção direcional irá promover ou não a diferenciação das populações, isso depende do ambiente e de eventos casuais. Duas populações que vivem em ambientes muito similares podem ser mantidas geneticamente similares por seleção direcional, mas, se existirem diferenças ambientais, a seleção pode dirigir as populações para composições diferentes.

A seleção que favorece os heterozigotos (**seleção balanceadora**) irá, na maior parte, manter polimorfismos mais ou menos similares em populações diferentes. Entretanto, novamente, se os ambientes são suficientemente diferentes, então as populações apresentarão alguma divergência. O oposto da seleção balanceadora é a seleção contra os heterozigotos, que produz equilíbrio instável. Tal seleção causará homozigose e divergência entre populações.

► 19.3 Vários Picos Adaptativos

Devemos evitar uma visão muito simplificada das consequências da seleção. Ao nível do gene, ou mesmo ao nível de um fenótipo parcial, há mais de um resultado possível da seleção para uma característica em determinado ambiente. A seleção para alterar uma característica (digamos, aumentar o tamanho) pode ser bem-sucedida de vários modos. Em 1952, Forbes Robertson e Eric Reeve selecionaram bem-sucedidamente a mudança do

tamanho da asa em duas populações diferentes de *Drosophila*. Entretanto, em um caso, o *número* de células na asa mudou, enquanto, em outro caso, o *tamanho* das células da asa mudou. Foram selecionados dois genótipos diferentes, ambos causando uma mudança no tamanho da asa. O estado inicial da população ao final da seleção determinou qual dessas seleções ocorreu.

Um simples caso hipotético ilustra como a mesma seleção pode levar a resultados diferentes. Suponha que a variação de dois loci (geralmente existe muito mais) influencia uma característica e (em determinado ambiente) os fenótipos intermediários têm a maior adaptabilidade. (Por exemplo, os neonatos têm uma maior chance de sobreviver se não forem nem muito grandes nem muito pequenos.) Se os alelos atuam de um modo simples influenciando o fenótipo, então as três constituições genéticas AB/ab , Ab/Ab e aB/aB produzirão uma alta adaptabilidade porque todos serão de fenótipo intermediário. Por outro lado, uma adaptabilidade muito baixa irá caracterizar homozigotos duplos AB/AB e ab/ab . Qual será o resultado da seleção? Podemos prever o resultado usando a adaptabilidade média \bar{W} de uma população. Como previamente discutido no Capítulo 17, a seleção atua nos casos mais simples para aumentar \bar{W} . Portanto, se calcularmos \bar{W} para cada combinação possível de frequências gênicas nos dois loci, podemos determinar que combinações produzem os valores altos de \bar{W} . Então devemos ser capazes de prever o curso da seleção seguindo uma curva de aumento de \bar{W} .

A superfície da adaptabilidade média para todas as possíveis combinações de frequências alélicas é chamada de **superfície adaptativa** ou **panorama adaptativo**, como ilustrada na Figura 19.5, que é como um mapa topográfico. A frequência do alelo A em um locus é plotada em um eixo, e a frequência do alelo B no outro locus é plotada no outro eixo. A altura acima do plano (representada por linhas topográficas) é o valor de \bar{W} que a população teria para uma combinação particular de frequências de A e B . De acordo com a regra de aumento de adaptabilidade, a seleção deverá levar a população de um “vale” de baixa adaptabilidade para um “pico” de alta adaptabilidade. Entretanto, a Figura 19.5 mostra que existem dois picos adaptativos, correspondendo a uma população fixa de Ab/Ab e uma população fixa de aB/aB , com um vale adaptativo entre eles. A qual pico a população irá ascender, e, portanto, qual é a composição final, depende de qual a composição genética inicial da população está de um lado ou outro da “linha tracejada” mostrada na Figura 19.5.

Sob condições idênticas de seleção natural, duas populações podem chegar a duas composições genéticas diferentes como um resultado direto da seleção natural.

É importante notar que nada na teoria da seleção requer que diferentes picos adaptativos sejam da mesma altura. A cinética da seleção dita apenas que \bar{W} aumenta, não que necessariamente atinja o pico mais alto possível no campo das frequências gênicas. Suponha, por exemplo, que uma população está perto do pico aB/aB na Figura 19.5 e que esse pico é mais baixo que o pico Ab/Ab . A seleção apenas não pode levar a população para Ab/Ab , pois isso iria requerer uma diminuição temporária em \bar{W} à medida que a população desceria de aB/aB , cruzaria a “sela” e ascenderia a outro topo. Assim, a força

Um panorama adaptativo mostra que a seleção tem múltiplos resultados possíveis

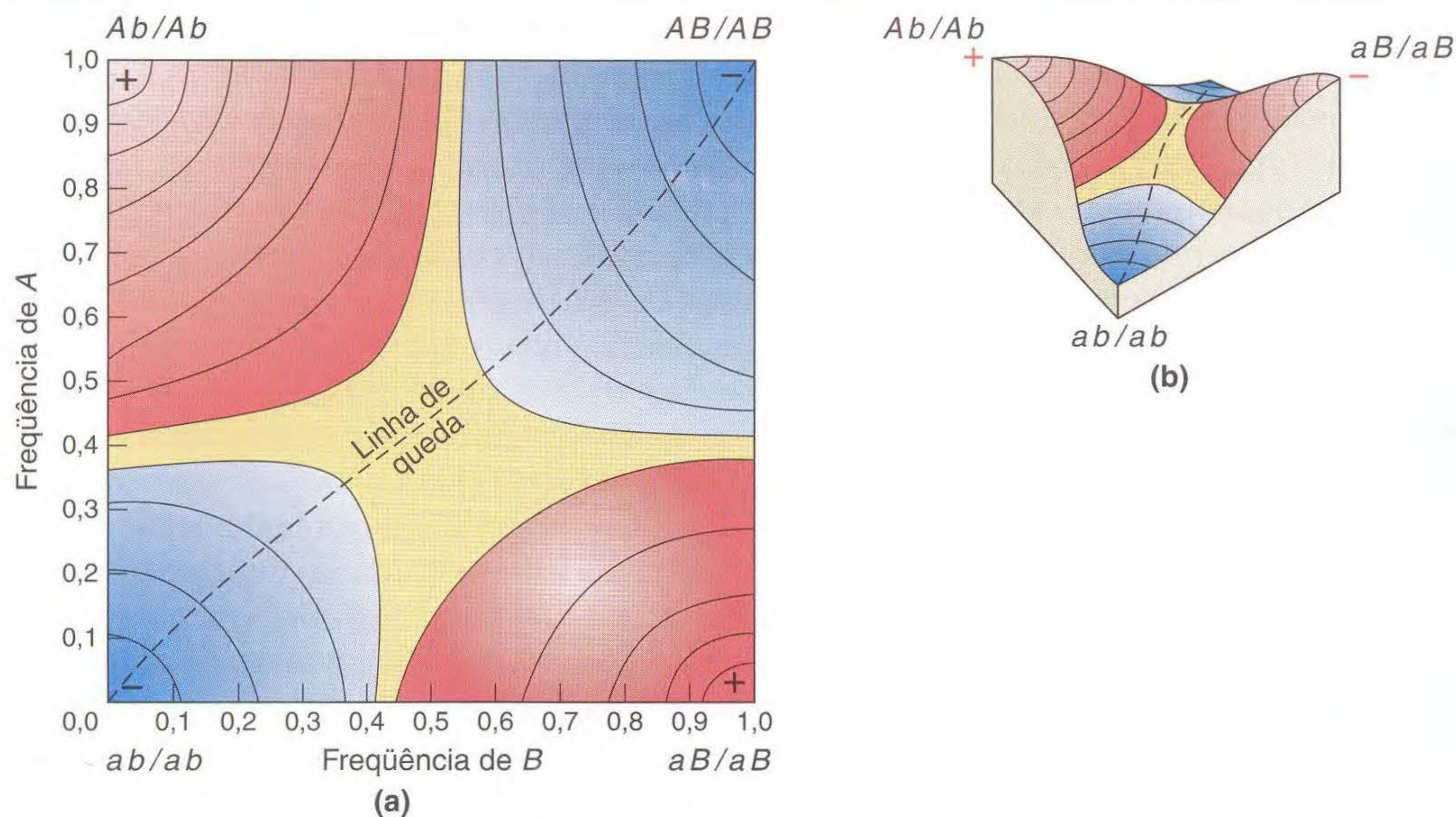


FIG. 19.5 Um panorama adaptativo com dois picos adaptativos (em vermelho), dois vales adaptativos (em azul) e uma sela topográfica no centro do panorama. As linhas topográficas são linhas de adaptabilidades médias iguais. Se a composição genética de uma população sempre muda de tal modo a mover a população "para cima" no panorama (aumento de adaptabilidade), então a composição final dependerá de onde a população começou com relação à linha de queda (tracejada). (a) O mapa topográfico do panorama adaptativo. (b) Uma perspectiva da superfície mostrando o mapa.

da seleção é míope. Ela dirige a população para um máximo *local* de W no campo das frequências gênicas, não a um *global*.

A existência de vários picos adaptativos para um processo seletivo significa que algumas diferenças entre as espécies são o resultado da história, e não das diferenças ambientais. Por exemplo, os rinocerontes africanos têm dois chifres no focinho, enquanto os rinocerontes indianos têm um. Não precisamos inventar uma história especial para explicar por que é melhor ter dois chifres nas planícies africanas e um na Índia. É muito mais plausível que a característica de ter dois chifres tenha sido selecionada, mas que dois chifres longos e finos e um curto, robusto, sejam simples características alternativas adaptativas, e as diferenças entre elas sejam o resultado de um acidente histórico. As explicações das adaptações por seleção natural não requerem que todas as diferenças entre as espécies sejam adaptativas.

Exploração dos picos adaptativos

As forças aleatórias e seletivas não devem ser imaginadas como simplesmente antagônicas. A deriva genética pode contrabalançar a força da seleção, mas também pode acentuá-la. O resultado do processo evolutivo é um resultado da operação simultânea dessas duas forças. A Figura 19.6 ilustra essas possibilidades. Note que existem muitos picos adaptativos nesse panorama. Devido à deriva aleatória, uma população sob seleção não sobe lentamente a um pico adaptativo. Ela toma um curso errático no campo das frequências gênicas, como um alpinista que fica sem oxigênio. A via I mostra uma história populacional onde a adaptação falhou. As flutuações aleatórias da frequência gênica foram suficientemente grandes para que a população por acaso ficasse fixada em um genótipo não-adaptável. Em qualquer população, alguma proporção de loci é fixada em um alelo seletivamente desfavorável porque a intensidade da seleção é insuficiente para superar a deriva

aleatória à fixação. A existência de vários picos adaptativos e a fixação aleatória de alelos menos adaptativos são características integrais do processo evolutivo. A seleção natural não pode basear-se na produção do melhor de todos os mundos possíveis.

A deriva genética pode alterar o resultado da seleção natural

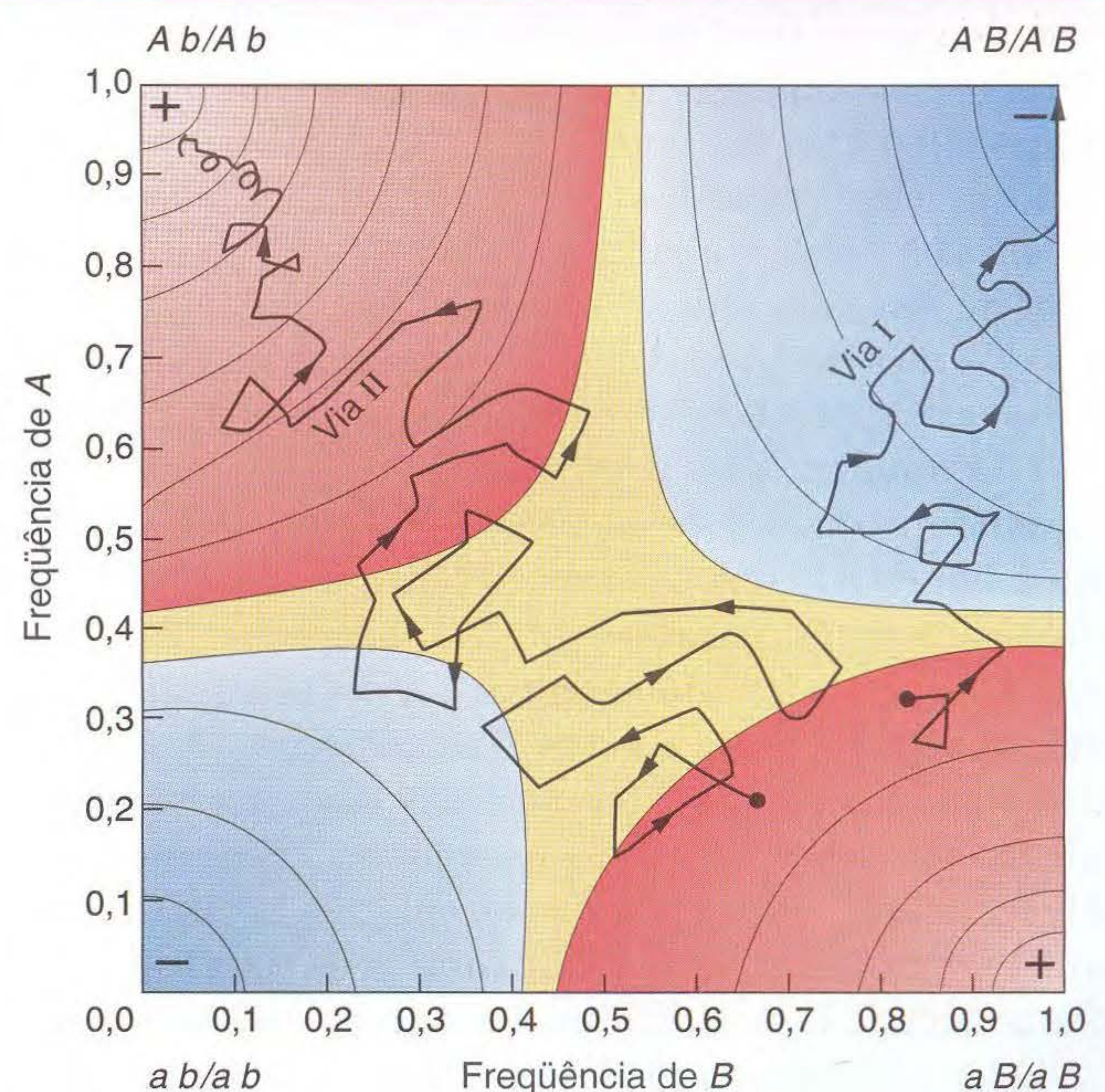


FIG. 19.6 A seleção e a deriva aleatória interagem para produzir mudanças diferentes na frequência gênica em um panorama adaptativo. Sem deriva genética, ambas as populações teriam se movido para aB/aB como resultado apenas de seleção.

A via II na Figura 19.6, por outro lado, mostra como a deriva aleatória pode melhorar a adaptação. A população originalmente estava na esfera de influência do menor pico adaptativo; entretanto, por flutuação aleatória em frequência gênica, sua composição passou pela “sela” adaptativa, e a população foi capturada pelo mais íngreme pico adaptativo. Essa passagem de um baixo estado adaptativo estável para outro mais alto poderia nunca ter ocorrido por seleção em uma população infinita, porque, apenas pela seleção, W nunca poderia diminuir temporariamente para cruzar de um lado para outro.

Outra fonte importante de indeterminação no resultado de um longo processo seletivo é a aleatoriedade no processo mutacional. Após as variações genéticas iniciais serem exauridas pela fixação seletiva e aleatória de alelos, nova variação surgindo de mutação pode ser a fonte de mais mudanças evolutivas. A direção particular dessa posterior evolução depende de mutações particulares que ocorrem e da ordem temporal na qual elas ocorrem. Uma ilustração muito clara dessa contingência histórica do processo evolucionário é um experimento de seleção feito por Holly Wichman e colaboradores. Eles forçaram o bacteriófago ϕ X174 a se reproduzir em altas temperaturas e no hospedeiro *Salmonella typhimurium* em vez de no seu hospedeiro normal *Escherichia coli*. Foram estabelecidas duas linhagens independentes de seleção, chamadas de TX e ID, e ambas desenvolveram a habilidade de se reproduzir em altas temperaturas no novo hospedeiro. Em uma das duas linhagens, a habilidade para se reproduzir em *E. coli* existia, mas, na outra linhagem, essa habilidade havia sido perdida. O bacteriófago tinha apenas 11 genes, e, assim, os experimentos foram capazes de registrar as mudanças sucessivas no DNA para todos esses genes e nas proteínas codificadas por eles durante o processo de seleção. Havia 15 mudanças de DNA na linhagem TX, situadas em 6 genes diferentes; na linhagem ID, havia 14 mudanças situadas em 4 genes diferentes. Em apenas 7 casos, as mudanças nas duas linhagens eram idênticas, incluindo uma grande deleção, mas mesmo essas mudanças idênticas apareciam em cada linhagem em uma ordem diferente. Assim, por exemplo, a mudança no sítio 1.533 do DNA, causando uma substituição de isoleucina por treonina, foi a 3ª mudança na linhagem ID, mas a 14ª mudança na linhagem TX.

Quando as mutações surgem em vários sítios na evolução de um estado fenotípico para outro, existem várias vias pelo espaço genético que a evolução pode tomar. Assim, se a diferença entre o fenótipo original e a forma evoluída é uma consequência de mutações em cinco sítios, A, B, C, D, e E, existem muitas ordens diferentes nas quais essas mutações poderiam ter ocorrido no tempo evolutivo. Primeiro, o sítio A pode ter sido fixado na população, então D, então C, então E e, finalmente, B. Por outro lado, a ordem da fixação evolutiva pode ter sido E, D, A, B, C. Para cinco sítios existem $5 \times 4 \times 3 \times 2 \times 1 = 120$ ordens possíveis. Duas perguntas importantes na compreensão da evolução são: Quantas dessas vias evolutivas alternativas são possíveis? e Quais as probabilidades das diferentes vias possíveis em relação umas as outras?

Daniel Weinreich e seus colaboradores caracterizaram em detalhes tal conjunto de vias evolutivas pelo espaço genético

em seus estudos da evolução da resistência a antibióticos na bactéria *E. coli*. A resistência ao antibiótico β -lactam cefotaxime é adquirida pelo acúmulo de cinco mutações em diferentes sítios no gene bacteriano de β -lactamase. Quatro das mutações levam a mudanças de aminoácidos, e a quinta é uma mutação não-codificante. Quando todas as cinco mutações estão presentes, a concentração mínima de antibiótico necessária para inibir o crescimento bacteriano aumenta por um fator de 100.000. Os pesquisadores primeiro dosaram a resistência conferida por uma mutação em determinado sítio na presença de todas as $2^4 = 16$ combinações possíveis de mutantes e não-mutantes nos outros quatro sítios. Em muitas combinações, mas não em todas, um mutante em um sítio era mais resistente, independentemente do estado dos outros quatro sítios. Por exemplo, um mutante no sítio G2385 mostrou uma resistência significativa, independentemente do estado mutante ou não-mutante dos outros quatro sítios. Por outro lado, a mutação do sítio não-codificante confere resistência significativa em oito combinações, mudança desprezível na resistência em seis casos, e uma *diminuição* na resistência em duas combinações. Essa dependência na vantagem ou desvantagem adaptativa de uma mutação nova nas mutações que foram previamente fixadas é o que os pesquisadores chamam de **epistasia de sinal**. Em seguida, eles dosaram a resistência em cada estágio na sequência temporal de adicionar mutações um sítio após o outro. Existem $5 \times 4 \times 3 \times 2 \times 1 = 120$ ordens possíveis nas quais os cinco sítios de mutação podem ocorrer. Se uma mutação em uma dessas possíveis ordens não confere uma alta resistência, então supostamente esse caminho evolutivo terminaria, pois não haveria seleção a favor da mutação nem mesmo contra ela. Eles descobriram que, das 120 possíveis vias na história mutacional, apenas 18 deram resistência aumentada em cada etapa mutacional. Assim, $102/120 = 85\%$ das possíveis vias mutacionais não eram acessíveis à evolução por seleção natural. Finalmente, supondo-se uma população que desenvolve resistência, a probabilidade de que uma via acessível particular seria de fato seguida é proporcional à magnitude da resistência aumentada em cada etapa, e, então, apenas 10 das 18 vias acessíveis corresponderiam a 90% dos casos de evolução de resistência bacteriana a antibióticos β -lactamais como a penicilina.

A ordem de ocorrência das mutações é de importância crítica para determinar se a evolução por seleção natural irá ou não atingir o estado mais vantajoso. Como a ordem de ocorrência das mutações é aleatória, muitos fenótipos vantajosos podem nunca ser atingidos, muito embora ocorram as mutações individuais.

► 19.4 Variação Genética

Herdabilidade da variação

A primeira regra de qualquer reconstrução ou previsão de mudança evolutiva é que a variação fenotípica seja herdável. É fácil construir histórias sobre a vantagem seletiva de uma

forma de característica sobre a outra, mas é uma questão de considerável dificuldade experimental mostrar que a variação na característica corresponde a diferenças genotípicas (veja o Capítulo 18). Não devemos supor que todas as características variáveis sejam herdáveis. Algumas características metabólicas (tais como resistência a altas concentrações de sal em *Drosophila*) mostram variação individual mas não herdabilidade. Em geral, as características comportamentais têm menores herdabilidades que as características morfológicas, especialmente em organismos com sistemas nervosos mais complexos que exibem imensa flexibilidade individual no sistema nervoso central. Antes que seja feita qualquer previsão sobre a evolução de determinada característica quantitativa, é essencial determinar se há uma variação genética para ela na população cuja evolução é prevista. Assim, as sugestões de que certas características na espécie humana, tais como desempenho de testes de QI, temperamento e organização social, estão no processo de evoluir ou ter evoluído em determinadas épocas na história humana dependem criticamente da evidência sobre se há variação genética para essas características. Reciprocamente, características que parecem ser fenotipicamente totalmente invariantes em uma espécie podem, entretanto, evoluir.

Um dos achados mais importantes na genética evolutiva foi a descoberta de que uma substancial variação genética pode ser subjacente a características que não apresentam variação morfológica. Elas são chamadas de **características canalizadas**, porque o resultado final de seu desenvolvimento é feito em limites estreitos a despeito de forças perturbadoras. Genótipos diferentes para características canalizadas têm o mesmo fenótipo constante na gama de ambientes que é usual para a espécie. As diferenças genéticas são reveladas se os organismos são postos em um ambiente estressante ou se uma grave mutação estressa o sistema desenvolvimental. Por exemplo, todas as *Drosophila* tipo selvagem têm exata-

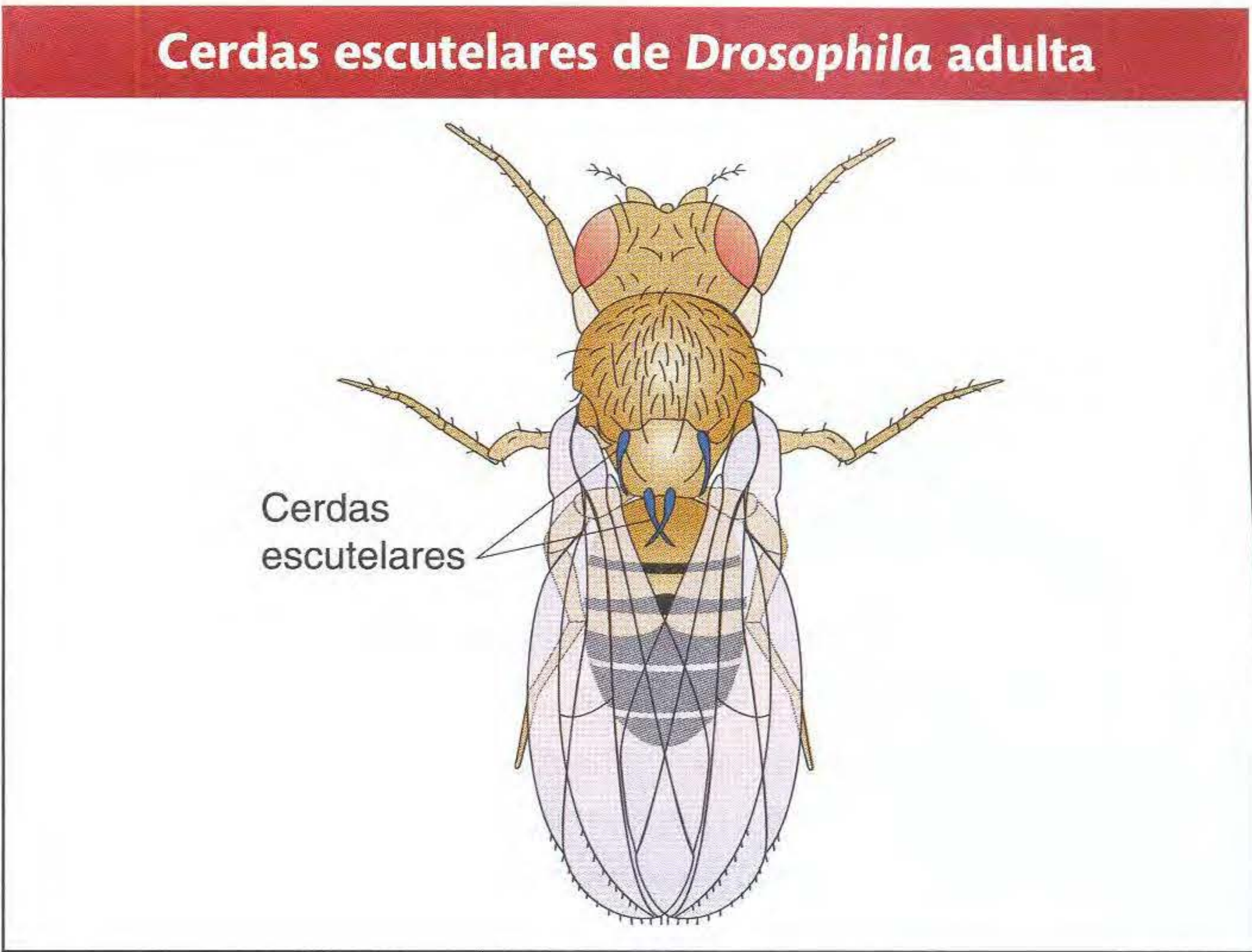


FIG. 19.7 As cerdas escutulares (mostradas em azul) são um exemplo de uma característica canalizada. Todas as *Drosophila* tipo selvagem têm quatro cerdas escutulares em uma ampla gama de ambientes.

mente quatro cerdas escutulares (Figura 19.7). Se o mutante recessivo *scute* estiver presente, o número de cerdas é reduzido, mas, em adição, há variação de mosca a mosca. Essa variação é herdável, e linhagens com zero ou uma cerda e linhagens com três ou quatro cerdas podem ser obtidas por seleção de moscas portadoras da mutação *scute*. Quando a mutação é removida, essas linhagens agora têm duas e seis cerdas, respectivamente. Experimentos similares com resultados similares foram feitos usando ambientes extremamente estressantes em lugar de mutantes. Uma consequência de tal variação genética oculta é que a característica que é fenotipicamente uniforme em uma espécie pode, entretanto, sofrer rápida evolução se um ambiente estressante revelar a variação genética.

Quadro 19.2 Exemplos de Diferenciação Extrema em Frequências de Grupos Sangüíneos Alélicos em Três Grandes Grupos Geográficos

Gene	Alelo	Grupo geográfico		
		Caucasóide	Negróide	Mongol
Duffy	<i>Fy</i>	0,0300	0,9393	0,0985
	<i>Fy^a</i>	0,4208	0,0000	0,9015
	<i>Fy^b</i>	0,5492	0,0607	0,0000
Rhesus	<i>R₀</i>	0,0186	0,7395	0,0409
	<i>R₁</i>	0,4036	0,0256	0,7591
	<i>R₂</i>	0,1670	0,0427	0,1951
	<i>r</i>	0,3820	0,1184	0,0049
	<i>r'</i>	0,0049	0,0707	0,0000
	Outros	0,0239	0,0021	0,0000
Antígeno P	<i>P₁</i>	0,5161	0,8911	0,1677
	<i>P₂</i>	0,4839	0,1089	0,8323

Fonte: Um resumo é dado em L.L. Cavalli-Sforza and W.F. Bodmer, *The Genetics of Human Populations* (W.H. Freeman and Company, 1971), pp. 724-731. Veja L.L. Cavalli-Sforza, P. Menozzi, and A. Piazza. *The History and Geography of Human Genes* (Princeton University Press, 1994), para dados detalhados.

Variação dentro de populações e entre populações

A variação genética existe entre populações aos níveis de morfologia, cariótipo, proteínas e DNA, como descrito no Capítulo 17. Vários desses exemplos também documentam algumas diferenças nas frequências genotípicas entre populações (veja Quadros 17.1 a 17.3 e 17.5). As quantidades relativas de variação dentro de populações e entre populações variam de espécie a espécie, dependendo da história e ambiente.

Em humanos, algumas frequências gênicas (por exemplo, as de cor da pele ou forma do cabelo) são bem diferenciadas entre populações e importantes grupos geográficos (chamadas raças geográficas). Se, entretanto, procuramos genes estruturais únicos, identificados imunologicamente ou por eletroforese e não por essas características fenotípicas, a situação é bem diferente. O Quadro 19.2 mostra os três loci para os quais caucasóides, negróides e mongóis são conhecidos como sendo muito diferentes uns dos outros (os grupos sanguíneos Duffy e Rhesus e o antígeno P). Mesmo para os loci mais divergentes, nenhum importante grupo geográfico é homozigoto para um alelo que está ausente nos outros dois grupos.

Em geral, populações humanas diferentes apresentam frequências bem similares para genes polimórficos. Os achados dos estudos de grupos sanguíneos polimórficos, loci codificantes de enzimas e polimorfismos de DNA em uma variedade de populações humanas mostraram que cerca de 85% da diversidade genética humana total são encontrados dentro de populações locais, cerca de 6% encontrados entre populações locais dentro de importantes raças geográficas, e os restantes 9%, encontrados entre as principais raças geográficas. Claramente, os genes que influenciam a cor da pele, forma do cabelo e forma facial que são bem diferenciadas entre as “raças” não são uma amostra aleatória de genes estruturais.

► 19.5 Mutação e Evolução Molecular

Embora a variação genética exista dentro das populações aos níveis de morfologia, cariótipo e proteínas, ela é, ao final, uma questão de variação nas seqüências de DNA. Cerca de um terço de todos os loci codificantes de proteínas são polimórficos, e todas as classes de DNA, incluindo éxons, íntrons, seqüências regulatórias e seqüências flangeadoras, mostram diversidade de nucleotídeos entre indivíduos dentro de populações.

As mutações de DNA podem ter três efeitos diferentes na adaptabilidade. Primeiro, elas podem ser deletérias, reduzindo a probabilidade de sobrevivência e reprodução de seus portadores. Todos os mutantes de laboratório usados na genética experimental têm algum efeito deletério na adaptabilidade. Segundo, elas podem aumentar a adaptabilidade aumentando a eficiência ou expandindo a faixa de condições ambientais nas quais a espécie pode viver ou permitindo que o organismo siga mudanças no ambien-

te. Terceiro, elas podem não ter efeito na adaptabilidade, deixando a probabilidade de sobrevivência e reprodução inalterada; são as chamadas mutações neutras. Para fins de compreensão da taxa de evolução molecular, entretanto, precisamos fazer uma distinção um pouco diferente, entre as mutações *efetivamente neutras* e *efetivamente selecionadas*. É possível provar que, em uma população finita de N indivíduos, o processo de deriva genética aleatória não será materialmente alterado se a intensidade da seleção, s , em um alelo é de ordem menor que $1/N$. Assim, a classe de mutações evolucionariamente neutras inclui tanto as que não têm nenhum efeito na adaptabilidade quanto aquelas cujos efeitos na adaptabilidade são menores que o recíproco do tamanho da população, tão pequenos que são efetivamente neutros. Por outro lado, se a intensidade da seleção, s , é de ordem maior do que $1/N$, então a mutação será efetivamente selecionada.

Gostaríamos de saber o quanto da evolução molecular é uma consequência de novas mutações adaptativas favoráveis que surgem nas espécies, o quadro apresentado por uma visão darwiniana simplista da evolução, e o quanto é simplesmente o acúmulo de mutações efetivamente neutras por fixação aleatória. Mutações que são efetivamente deletérias não precisam ser consideradas, pois serão mantidas em baixas frequências nas populações e não contribuem para mudança evolutiva. Se uma mutação recém-surgida é efetivamente neutra, então, como indicado no Capítulo 17, há uma probabilidade de $1/(2N)$ de que ela substitua o alelo prévio devido à deriva genética aleatória. Se μ é a taxa de aparecimento de mutações efetivamente neutras em um locus por cópia de gene por geração, então o número absoluto de novas cópias mutacionais que aparecerão em uma população de N indivíduos diplóides é $2N\mu$. Cada uma dessas novas cópias tem a probabilidade de $1/(2N)$ de eventualmente se instalar na população. Assim, a taxa absoluta de substituição dos alelos antigos por novos em um locus por geração é sua taxa de surgimento multiplicada pela possibilidade de que qualquer uma delas ocorra por deriva:

$$\text{taxa de substituição neutra} = 2N\mu \times 1/(2N) = \mu$$

Ou seja, esperamos que, a cada geração, existam μ substituições de um novo alelo por um antigo em cada locus na população, puramente por deriva genética de mutações efetivamente neutras.

A taxa de substituição na evolução resultante de deriva genética aleatória de mutações efetivamente neutras é igual à taxa de mutação de tais alelos, μ .

Sinal característico da seleção purificadora no DNA

A taxa constante de substituições neutras prevê que, se o número de diferenças de nucleotídeos entre duas espécies fosse plotado contra o tempo desde sua divergência de um ancestral comum, o resultado seria uma linha reta com uma

A taxa de mutação é mais alta em sítios sinônimos do que nos sítios não-sinônimos

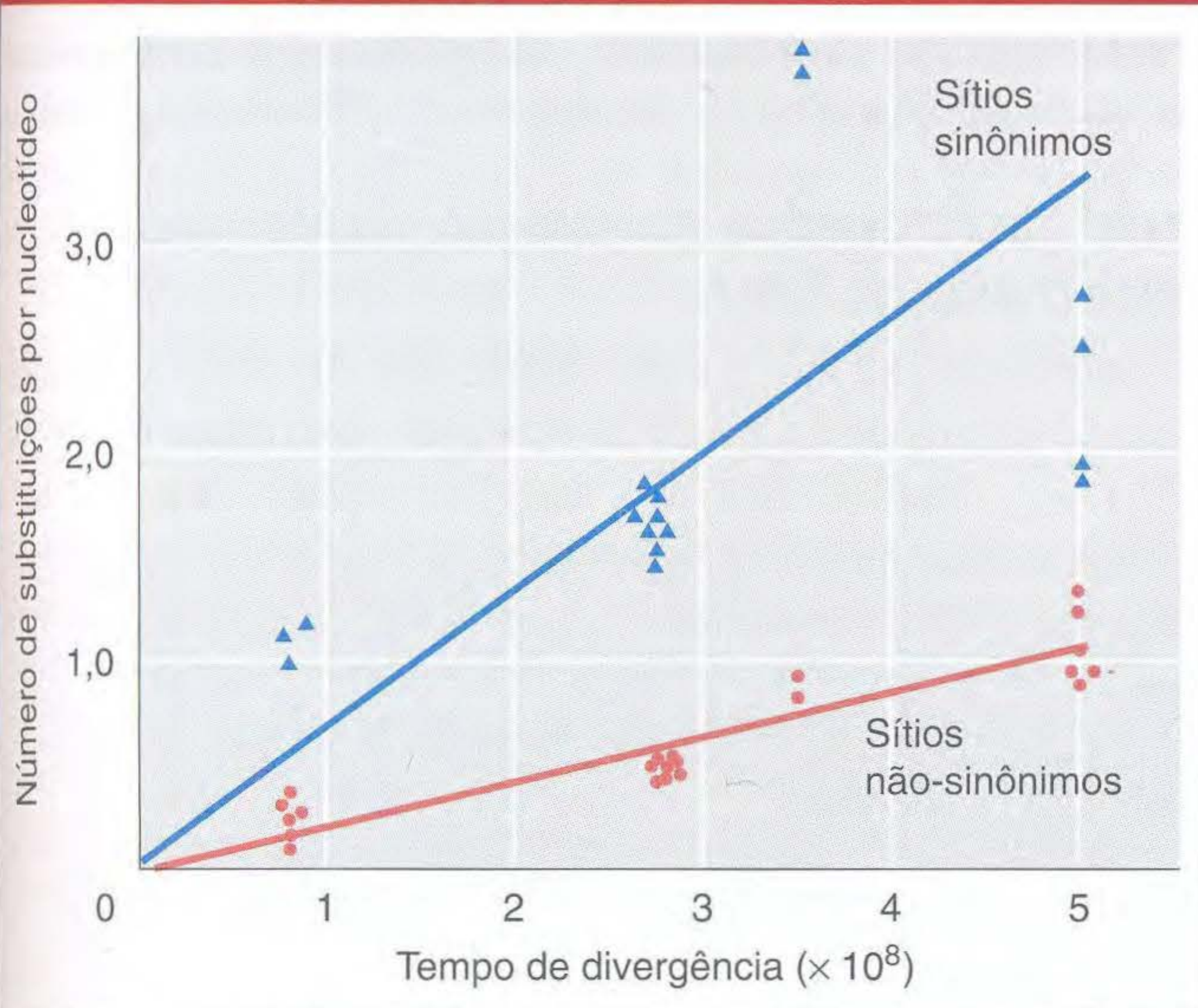


FIG. 19.8 A quantidade de divergência de nucleotídeos em sítios sinônimos é maior do que a quantidade de divergência em sítios não-sinônimos do gene de β -globina.

variações danosas”, como Darwin chamava, é difusa. A seleção purificadora explica por que encontramos muitas seqüências de proteínas que estão inalteradas ou quase inalteradas durante o longo tempo evolutivo. Por exemplo, um conjunto de cerca de 500 genes que existem em todos os domínios da vida — Archaea, bactéria, fungos, plantas e animais — codificam proteínas cujas seqüências foram amplamente conservadas durante mais de 3 bilhões de anos de evolução. Para preservar tais seqüências, as variantes que surgiram aleatoriamente em indivíduos de dezenas de milhões de espécies foram rejeitadas repetidamente pela seleção.

A seleção purificadora é um aspecto difuso da seleção natural que reduz a variação genética e preserva as seqüências do DNA e proteínas durante muito tempo.

Outra previsão da evolução neutra é que proteínas diferentes terão velocidades (*clock rates*) diferentes, porque as funções metabólicas de algumas proteínas serão muito mais sensíveis a mudanças em suas seqüências de aminoácidos. As proteínas nas quais cada aminoácido faz uma diferença terão uma taxa de mutação neutra efetivamente menor devido à menor proporção de suas mutações que serão neutras comparadas a proteínas que são mais tolerantes à substituição. A Figura 19.9 mostra uma comparação dos relógios para fibrinopeptídeos, hemoglobina e citocromo *c*. Que os fibrinopeptídeos tenham uma proporção muito maior de mutações neutras é razoável porque esses peptídeos são meramente uma segurança não-metabólica, removida do fibrinogênio para ativar a reação de coagulação sangüínea. De considerações *a priori*, o motivo pelo qual as hemoglobinas são menos sensíveis a mudanças de aminoácidos do que o citocromo *c* é menos óbvio.

As proteínas diferem na taxa de mutação

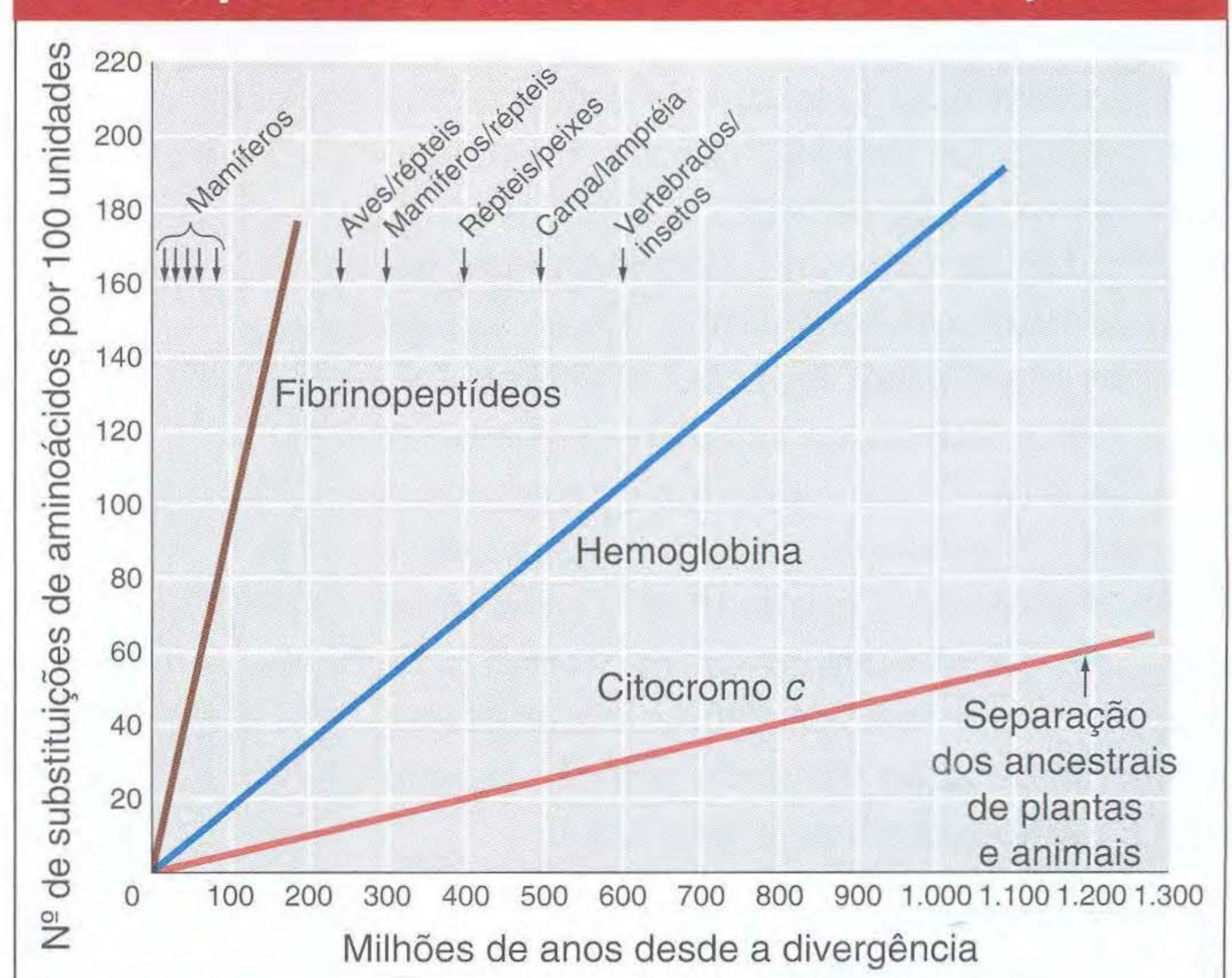


FIG. 19.9 Número de substituições de aminoácidos na evolução dos vertebrados como função do tempo desde a divergência. As três proteínas, fibrinopeptídeos, hemoglobina e citocromo *c*, diferem na taxa de substituição porque as proporções diferentes de suas substituições de aminoácidos são seletivamente neutras.

inclinação igual a μ . Isto é, a evolução deveria continuar de acordo com um **relógio molecular** que está batendo a uma velocidade μ . A Figura 19.8 mostra tal plotagem para o gene de β -globina. Os resultados são bem consistentes com a indicação de que as substituições de nucleotídeos foram efetivamente neutras nos últimos 500 milhões de anos. Dois tipos de substituições de nucleotídeos são plotados: **substituições sinônimas**, que são de um códon alternativo para outro, não alterando o aminoácido, e **substituições não-sinônimas**, que resultam em uma mudança de aminoácido. A Figura 19.8 mostra uma inclinação muito menor para substituições não-sinônimas do que as mudanças sinônimas, o que significa que a taxa de mutação para substituições não-sinônimas seletivamente neutras é mais baixa do que as sinônimas.

Esse resultado é exatamente o que se espera sob seleção natural. As mutações que causam uma substituição de aminoácidos devem ter um efeito deletério acima do limiar de evolução neutra mais frequentemente que as substituições sinônimas, que não mudam a proteína. Tais variantes deletérias serão removidas da seleção pela **seleção purificadora**. Uma taxa menor que a esperada de não-sinônimas para alterações sinônimas é um sinal característico de seleção purificadora. É importante notar que essas observações não mostram que as substituições sinônimas não têm restrições seletivas nelas; elas mostram que essas restrições são, em média, não tão fortes quanto as mutações que mudam os aminoácidos. Assim, uma mudança sinônima, embora não tenha efeito na seqüência de aminoácidos, muda o mRNA para essa seqüência e, assim, pode afetar a estabilidade, o tempo de vida e a velocidade de tradução da mensagem.

A seleção purificadora é a faceta mais generalizada, mas frequentemente esquecida da seleção natural. A “rejeição de

A taxa de evolução neutra para a sequência de aminoácidos de uma proteína depende da sensibilidade do funcionamento da proteína a mudanças de aminoácidos.

Como as sequências de DNA e proteínas mudam em função do tempo, decifrar que mudanças são adaptativas e quais são neutras tem sido um desafio considerável. Nas três seções seguintes do capítulo, ilustraremos alguns modos e exemplos nos quais as mudanças genéticas foram ligadas à diversidade organismal.

► 19.6 Relação entre Genética e Mudança Funcional na Evolução de Proteínas

Como tanta evolução de sequência é efetivamente neutra, não há relação simples entre a quantidade de mudança no DNA de um gene e a quantidade de mudança, se houver alguma, na função codificada da proteína. Em um extremo, quase toda a sequência de aminoácidos de uma proteína pode ser substituída, embora mantenha a função original. Por exemplo, os eucariontes, de leveduras a humanos produzem lisozimas, uma enzima que degrada as paredes das células bacterianas. Praticamente cada aminoácido nessa proteína foi substituído desde que as linhagens de leveduras e vertebrados divergiram de um ancestral comum. Assim, um alinhamento de suas duas proteínas ou sequências de DNA não revelaria nenhuma similaridade. A evidência de que genes de lisozimas de leveduras e humanas descendem de um gene ancestral original comum vem de comparações de formas evolutivas intermediárias que apresentam mais e mais divergências de sequências à medida que as espécies se tornam mais divergentes. A enzima manteve sua função a despeito da substituição de aminoácidos porque os aminoácidos certos foram substituídos para manter a estrutura tridimensional da enzima.

Em contraste, a função de uma enzima pode ser alterada por uma única substituição de aminoácidos. A mosca varejeira de ovelhas, *Lucilia cuprina*, desenvolveu resistência a inseticidas de organofosfato amplamente usados para controlá-la. Richard Newcombe, Peter Campbell e seus colaboradores mostraram que essa resistência é a consequência de uma única substituição de uma glicina por ácido aspártico no sítio ativo de uma enzima que geralmente é uma carboxilesterase. A mutação causa a perda completa da atividade de carboxilesterase e é substituída pela especificidade de esterase. O modelamento tridimensional da molécula indica que a proteína substituída ganha a habilidade de se ligar a uma molécula de água perto do sítio de ligação do organofosfato, que é então hidrolisado pela água.

Não há relação regular entre quanta mutação de DNA ocorre na evolução e quantas mudanças funcionais resultam.

O papel da seleção na evolução da carboxilesterase de insetos e na resistência a inseticidas é óbvio. Na maioria dos casos,

entretanto, os padrões de substituição de aminoácidos são mais complexos, e o papel da seleção na mudança de proteínas é menos aparente. Para detectar a seleção nesses casos, foram criados métodos para descobrir um sinal característico indicando que a seleção agiu nas sequências de DNA e de proteínas.

Sinal característico da seleção positiva nas sequências de DNA

A demonstração do relógio molecular diz que a maioria das substituições de nucleotídeos que ocorreram na evolução eram neutras, mas não nos diz o quanto da evolução molecular é adaptativa. Um modo para detectar a evolução adaptativa de uma proteína é comparar os polimorfismos sinônimos e não-sinônimos de nucleotídeos dentro de espécies com mudanças sinônimas e não-sinônimas de nucleotídeos entre espécies. No funcionamento da evolução neutra por deriva genética aleatória, o polimorfismo dentro de uma espécie é simplesmente um estágio na fixação eventual de um novo alelo; logo, se todas as mutações são neutras, a proporção de não-sinônimos para polimorfismos sinônimos de nucleotídeos dentro de uma espécie deve ser a mesma que a proporção de substituições de nucleotídeos não-sinônimos para sinônimos entre as espécies. Por outro lado, se as mudanças de aminoácidos entre espécies foram ativadas por uma seleção adaptativa positiva, deve haver um excesso de mudanças não-sinônimas entre as espécies. O Quadro 19.3 mostra uma aplicação desse princípio por John MacDonald e Martin Kreitman ao gene da álcool desidrogenase em três espécies bem próximas de *Drosophila*. Claramente, há um excesso de substituições de aminoácidos entre as espécies em relação ao esperado dos polimorfismos. Portanto, concluímos que algumas substituições de aminoácidos eram mudanças adaptativas ativadas pela seleção natural.

Evolução morfológica

Nos exemplos examinados até agora, as mudanças estruturais em enzimas alteraram a especificidade do substrato, e as consequências fenotípicas são limitadas ao metabolismo. Mudanças visíveis mais marcantes nos fenótipos também podem surgir por meio de mutações nas sequências codificantes de proteínas quando essas proteínas são membros de vias reguladoras que controlam a diferenciação em tipos celulares distintos no início do desenvolvimento dos organismos. Alguns dos exemplos mais marcantes e mais bem

Quadro 19.3 Polimorfismos Sinônimos e Não-sinônimos e Diferenças de Espécies para Álcool Desidrogenase em Três Espécies de *Drosophila*

	Diferenças de espécies	Polimorfismos
Não-sinônimas	7	2
Sinônimas	17	42
Proporção	0,29:0,71	0,05:0,95

Fonte: J. McDonald and M. Kreitman, "Adaptive Protein Evolution at the *Adh* locus in *Drosophila*," *Nature* 351, 1991, 652–654.

compreendidos da divergência fenotípica são encontrados nos padrões de cor do corpo dos animais. A pelagem dos mamíferos, plumagem das aves, escamas de peixes e cor das asas em insetos são maravilhosamente diversos. Muito progresso foi feito na compreensão do controle genético da formação da cor e seu padrão na evolução das diferenças de cor dentro das espécies e entre as espécies.

Na região de Pinacate, ao Sudoeste do Arizona, formações rochosas são circundadas por granito arenoso de cor mais cla-

ra (Figura 19.10). O pequeno camundongo das rochas, *Chaetodipus intermedius*, habita Pinacate bem como outras áreas rochosas do Sudoeste. Os camundongos encontrados nos depósitos de lava são tipicamente de cor escura, enquanto os encontrados nas áreas vizinhas de granito de cor arenosa ou no solo do deserto são geralmente de cor clara (Figura 19.11). Essas diferenças de cor sugerem uma preferência de hábitat que pode proteger os camundongos de serem vistos pelos predadores.

Cores contrastantes do deserto Pinacate



FIG. 19.10 Pedras de lava no deserto Pinacate produziram crescimentos de rochas de cor preta adjacentes a substratos de cor de areia. [Foto por cortesia de Michael Nachman, University of Arizona.]

As substituições de aminoácidos acumulam-se em uma parte da proteína MC1R



FIG. 19.11 *Chaetodipus intermedius* de cor clara e de cor escura da região de Pinacate do Arizona são mostrados sobre fundos de rochas de lava cor de areia e escura. [Fotos por cortesia de Michael Nachman, de M.W. Nachman, H.E. Hoekstra, and S.L. D'Agostino. "The Genetic Basis of Adaptive Melanism in Pocket Mice," *Proc. Natl. Acad. Sci. USA* 100, 2003, 5268-5273.]

O pequeno camundongo das rochas é um exemplo de **melanismo**, a ocorrência de uma forma escura dentro de uma população ou espécie. O melanismo é um dos tipos mais comuns de variação fenotípica em animais. A cor escura da pelagem é devida a uma intensa deposição do pigmento **melanina**, o pigmento mais comum no reino animal. Em mamíferos, dois tipos de melanina são produzidos nos melanócitos (as células pigmentadas da epiderme e folículos pilosos): a eumelanina, que forma pigmentos pretos ou marrons, e a feomelanina, que forma pigmentos amarelos ou vermelhos. As quantidades relativas de eumelanina e feomelanina são controladas pelos produtos de vários genes. Duas proteínas importantes são o receptor de melanocortina 1 (MC1R) e a proteína aguti. Durante o ciclo de crescimento do pêlo, o hormônio estimulante de α -melanócito (α -MSH) liga-se à proteína MC1R, que ativa a indução de enzimas produtoras de pigmento. A proteína aguti é um antagonista de ativação de MC1R e inibe a produção de eumelanina.

Michael Nachman e seus colaboradores examinaram as seqüências de DNA dos genes *mc1r* dos camundongos de cores clara e escura. Eles encontraram a presença de quatro mutações no gene *mc1r* no camundongo escuro que faz com que a proteína MC1R difira em quatro aminoácidos da proteína correspondente no camundongo claro. Os achados dos estudos bioquímicos sugerem que tais mutações fazem com que a proteína MC1R seja constitutivamente ativa (ativa o tempo todo), suplantando a regulação do receptor de atividade pela proteína aguti. De fato, as mutações em *mc1r* estão associadas ao melanismo em todos os tipos de vertebrados selvagens e domesticados. Muitas dessas mutações alteram aminoácidos na mesma parte da proteína MC1R, e as mesmas mutações ocorrem independentemente em algumas espécies (Figura 19.12).

De muitos modos, podemos pensar nesses camundongos escuros como análogos aos tentilhões de Darwin e a lava como novos habitats da “ilha”, sendo produzida pela mesma atividade vulcânica que produziu as ilhas Galápagos. A forma com cor de areia do camundongo parece ser o tipo ancestral, similar ao tentilhão ancestral continental que colonizou Galápagos. Se a vantagem de ser menos visível aos predadores resultou em seleção natural para cor da pelagem, a invasão das ilhas de pedras de lava pelos camundongos levou à dispersão de um alelo que era favorecido no fundo preto das rochas e selecionado contra no fundo cor de areia. Novas mutações no gene *mc1r* seriam essenciais para tal mudança adaptativa.

Inativação gênica

Há muito foi notado que os animais que vivem em cavernas geralmente são cegos e descoloridos. Darwin escreveu no *The Origin of Species* que “vários animais pertencentes a classes muito diferentes, que habitam as cavernas de Carniola [na Eslovênia] e Kentucky, são cegos. Como é difícil imaginar que os olhos, embora inúteis, possam de algum modo ser prejudiciais aos animais que vivem no escuro, sua perda pode ser atribuída ao desuso.”

Muitas espécies de peixes que vivem em cavernas perderam seus olhos e a cor do corpo. Como essas espécies per-

As substituições de aminoácidos acumulam-se em uma parte da proteína MC1R

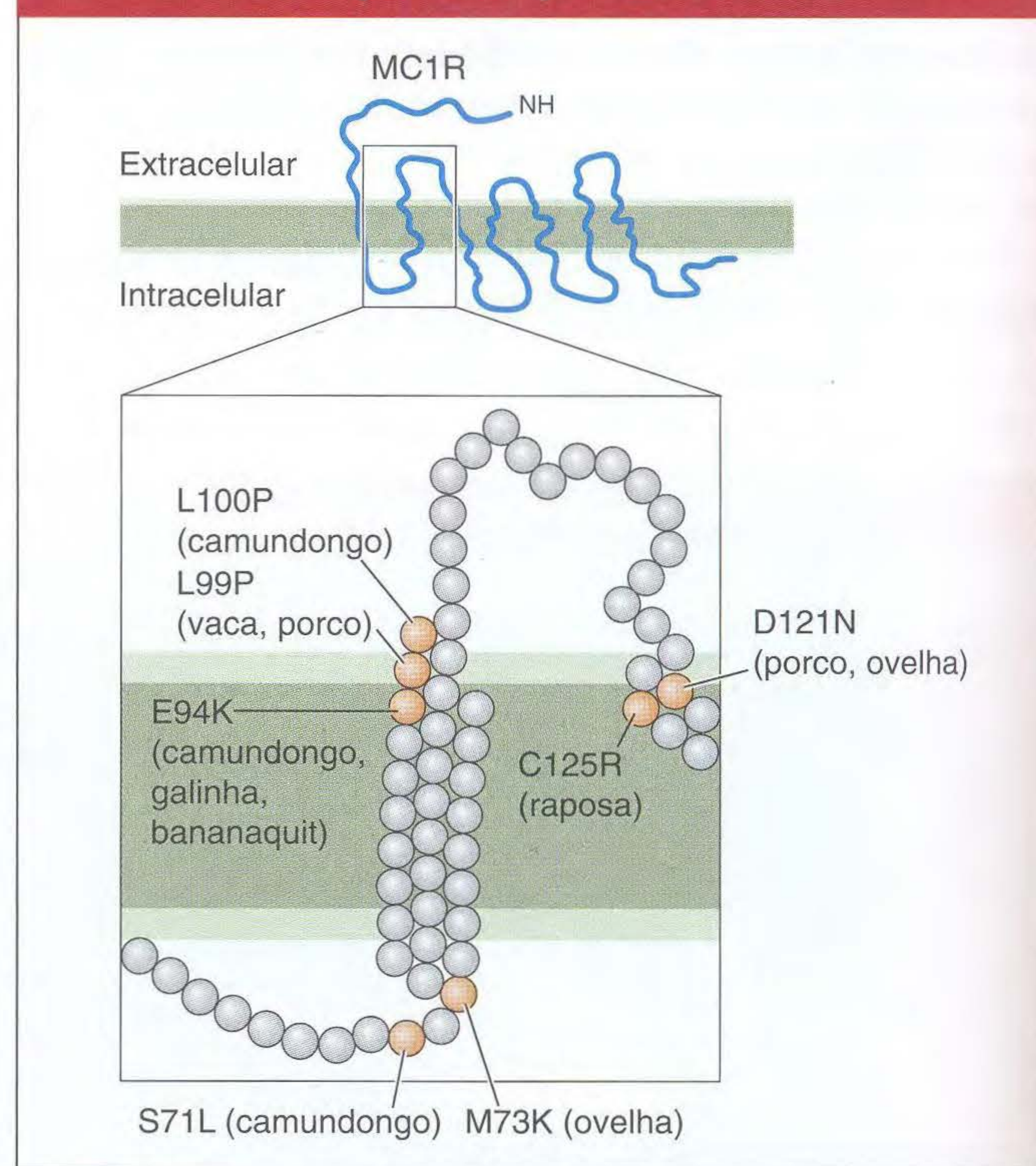


FIG. 19.12 Substituições de aminoácidos (círculos laranja) associadas a melanismo variam levemente de local em espécies diferentes, mas estão situadas na mesma parte da proteína MC1R. A parte superior da figura mostra a topologia geral da proteína MC1R. A região na qual estão situadas as substituições está aumentada na parte inferior da figura. [De E. Eizirik et al. "Molecular Genetics and Evolution in the Cat Family," *Curr. Biol.* 13, 2003, 448-453; reimpressa com permissão de Elsevier.]

tencem a famílias diferentes que incluem viver na superfície, espécies com olhos, a perda dos olhos e da pigmentação claramente ocorreu repetidamente. Por exemplo, o peixe cego das cavernas mexicanas (*Astyonax mexicanus*) pertence à mesma ordem das piranhas e à tetra cor de neon. Cerca de 30 populações de caverna no México perderam a cor do corpo de seus parentes que vivem na superfície (Figura 19.13).

Os estudos genéticos indicam que o albinismo na população de peixes Pachón de caverna é devido a uma única mutação recessiva. Além disso, um cruzamento entre um indivíduo da caverna Molino e um indivíduo Pachón produziu apenas prole albina, sugerindo que o albinismo nas duas populações é devido ao mesmo locus genético. Para identificar o gene responsável pelo albinismo em peixe, os pesquisadores investigaram os genótipos de peixes em vários loci de pigmentação conhecidos como causando albinismo em camundongos ou humanos. Eles descobriram que um desses genes, *Oca2*, estava mapeado no locus albino, e que havia uma associação perfeita entre o genótipo do locus *Oca2* e o fenótipo de albinismo na prole F_2 que era um retrocruzamento entre a prole F_1 Molino e Molino/superfície ou F_1 Pachón e Pachón/superfície.

Evolução de albinismo em peixe cego de caverna



FIG. 19.13 Formas da superfície do peixe *Astyanu mexicanus* parecem normais, mas as populações de caverna, tais como as das cavernas de Molino e Pachón no México, estão repetidamente envolvidas em cegueira e albinismo. [Fotos por cortesia de Meredith Protas and Cliff Tabin, Harvard Medical School.]

Uma maior inspeção do gene *Oca2* revelou que a população Pachón era homozigota para uma deleção que ia desde um íntron através da maioria de um éxon, e que a população Molino era homozigota para uma deleção de um éxon diferente. As análises funcionais provaram que cada deleção no gene *Oca2* causava a perda da função de *Oca2*.

A identificação de lesões diferentes no gene *Oca2* das duas populações de caverna indica que o albinismo evoluiu separadamente nas duas populações de caverna. Existem também evidências de que uma terceira população de caverna tinha uma terceira mutação *Oca2* distinta. É conhecido de outros vertebrados que o albinismo pode evoluir por mutações em outros genes. O que pode explicar, então, a repetida inativação do gene *Oca2*?

Existem duas explicações possíveis. Primeiro, as mutações *Oca2* parecem não causar graves defeitos colaterais além da perda de pigmentação e visão. Alguns outros genes de pigmentação, quando mutados em peixes, causam marcantes reduções em viabilidade. Os efeitos das mutações *Oca2* parecem, assim, ser menos *pleiotrópicos* e ter efeitos na adaptabilidade geral que são menos prejudiciais que os de mutações em outros genes de pigmentação de peixes. Segundo, o locus *Oca2* é muito grande, cobrindo uns 345 kb em humanos e contendo 24 éxons. Ele representa um alvo muito grande para mutações aleatórias que perturbariam o funcionamento do gene; as mutações *Oca2*, portanto, são mais prováveis de surgir do que as mutações em pequenos loci.

A perda funcional de genes não é o que geralmente pensamos quando refletimos sobre evolução. Mas a inativação gênica é certamente o que devemos prever que aconteça

quando as condições seletivas mudam ou quando as populações de espécies mudam seus habitats ou estilos de vida e algumas funções gênicas não são mais necessárias. Por exemplo, a bactéria *E. coli* tem genes para as enzimas que são necessárias para produzir todos os aminoácidos essenciais; assim, ela pode crescer em um meio de cultura que contém apenas açúcar, uma fonte inorgânica de nitrogênio e uma vitamina. Os humanos, como outros animais, não têm esses genes e rapidamente acabariam morrendo em tal dieta.

► 19.7 Evolução Regulatória

Como já discutido, uma importante restrição à evolução gênica é o potencial efeito pleiotrópico de mutações nas regiões codificantes. Esses efeitos podem ser evitados por mutações nas seqüências regulatórias, que têm um papel importante na evolução da regulação gênica e forma do corpo.

Os exemplos de evolução na coloração do corpo discutidos até agora afetavam todo o padrão de pelagem ou escamas. A evolução do preto ou totalmente despigmentado na coloração do corpo pode surgir por mutações nas funções de genes de pigmentação. Entretanto, muitos esquemas de cor em geral são feitos de duas ou mais cores em algum padrão espacial. Em tais casos, a expressão de genes de pigmentação deve diferir em áreas do corpo que serão de cores diferentes. Em populações diferentes ou espécies, a regulação de genes de pigmentação devem evoluir por algum mecanismo que não perturba a função de proteína de pigmentação.

A espécie do gênero de mosca-das-frutas *Drosophila* apresenta uma ampla diversidade de marcas no corpo e asas. Um padrão comum é a presença de um ponto preto perto da ponta da asa nos machos (Figura 19.14). A produção dos pontos pretos requer enzimas que sintetizam melanina, o mesmo pigmento feito no pequeno camundongo. Muitos genes que

Manchas nas asas da mosca-das-frutas

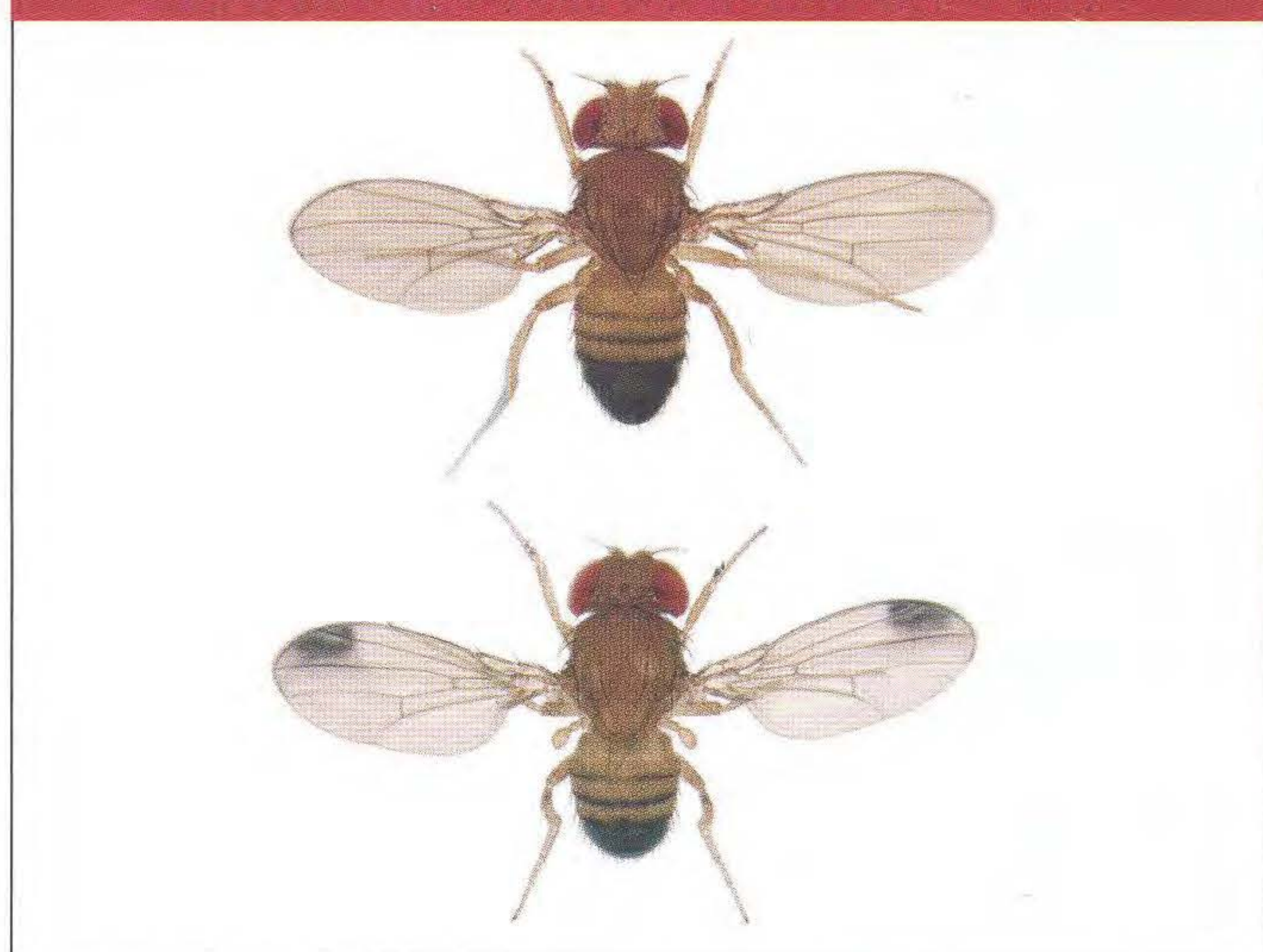


FIG. 19.14 Machos de *Drosophila melanogaster* não apresentam manchas na asa (em cima), enquanto os machos de *Drosophila biarmipes* (embaixo) têm manchas escuras nas asas, que são exibidas em um ritual de corte. Essa simples diferença morfológica é devida a diferenças na regulação de genes de pigmentação. [Fotos de Nicolas Gompel.]

controlam a via de síntese de melanina foram bem estudados no organismo-modelo *Drosophila melanogaster*. Um gene, chamado de *yellow* porque as mutações nele observadas causam áreas de pigmentação escura no corpo que parecem amareladas ou bronzeadas, tem um papel central no desenvolvimento e divergência dos padrões de melanina. Nas espécies com manchas, a proteína Yellow é expressa em altos níveis nas células das asas que produzirão o ponto preto, enquanto, na espécie sem pontos, Yellow é expressa em baixo nível nas asas (Figura 19.15a).

A diferença na expressão de Yellow entre as espécies com manchas e sem manchas pode ser devida a diferenças na localização espacial de fatores de transcrição que regulam *yellow* (i.e., mudanças em seqüências de ação trans do gene *yellow*) e/ou a diferenças nas seqüências reguladoras de ação cis que controlam como o gene *yellow* é regulado. Para examinar que mecanismos estão envolvidos, a atividade das seqüências reguladoras de ação cis de *yellow* de espécies diferentes foi examinada colocando-as antecedentes a um gene repórter e introduzindo-as em *D. melanogaster*.

O gene *yellow* é regulado por um arranjo de elementos regulatórios de ação cis que governam a transcrição gênica em tecidos diferentes e tipos celulares e épocas diferentes do desenvolvimento (Figura 19.15b). Esses elementos incluem os que controlam a transcrição das partes bucais larvares, o tórax pupal e o abdome, e o desenvolvimento das asas. Foi descoberto que, enquanto o elemento regulatório de ação cis da espécie sem manchas nas asas ativa um baixo nível de expressão de um gene repórter ao longo da asa, o elemento correspondente de uma espécie com manchas, tal como *D. biarmipes* ou *D. elegans*, ativa um alto nível de expressão de um gene repórter em uma mancha perto da ponta da asa (Figura 19.15c). Essas observações mostram que mudanças

na seqüência e função de um elemento regulatório de ação cis são responsáveis pela mudança na regulação de *yellow* e contribuem para a origem da mancha na asa. O elemento regulatório de ação cis da espécie com manchas aparentemente tinha adquirido sítios de ligação para fatores de transcrição que agora ativam altos níveis de transcrição gênica em um padrão de mancha na asa em desenvolvimento.

O papel crítico da seqüência regulatória de ação cis na evolução da regulação gênica e características morfológicas pode ser mais bem explicado à luz dos efeitos pleiotrópicos de mutações codificantes de genes de “ferramentas” (*toolkit*). Nesse caso, o gene *yellow* é altamente pleiotrópico: ele é necessário para a pigmentação de muitas estruturas e também em funções no sistema nervoso. Uma mutação codificante que afeta a atividade da proteína Yellow afetaria a atividade Yellow em *todos* os tecidos, o que pode ter consequências negativas para a adaptabilidade. Entretanto, como os elementos regulatórios individuais de ação cis geralmente afetam apenas um aspecto da expressão gênica, as mutações nas seqüências regulatórias de ação cis fornecem um mecanismo para mudar um aspecto da expressão gênica, preservando o papel de proteínas pleiotrópicas em outros processos desenvolvimentais.

Evitar o efeito pleiotrópico de mutações codificantes é certamente uma consideração muito importante para desenvolver novos papéis para fatores de transcrição que possam regular dúzias a centenas de genes-alvo. As mudanças nas seqüências codificantes de um fator de transcrição, por exemplo, o domínio de ligação ao DNA, podem afetar todos os genes-alvo com consequências catastróficas para o animal. Entretanto, a maioria dos fatores de transcrição, tais como os ferramentais (*toolkit*) animais, são regulados, como descrito no Capítulo 12, por vários elementos regulatórios de ação cis.

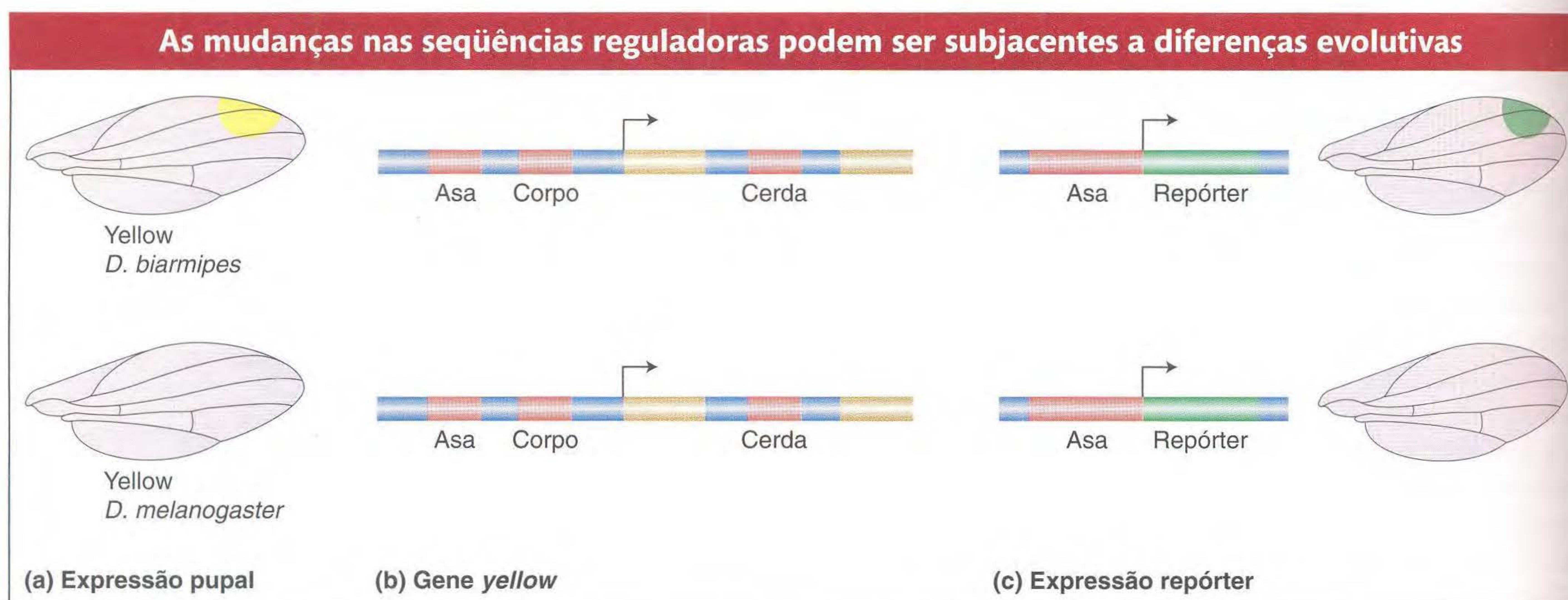


FIG. 19.15 A evolução da regulação do gene e morfologia no caso é devida à evolução em seqüências reguladoras de ação cis. (a) Na mosca-da-fruta com manchas, a proteína de pigmentação Yellow é expressa em altos níveis nas células que irão produzir grandes quantidades de melanina. (b) O locus *yellow* de *Drosophila* contém vários discretos elementos de ação cis (vermelho) que controlam a transcrição de *yellow* em partes diferentes do corpo. Os éxons são mostrados em amarelo. As setas indicam a regulação do éxon. (c) O elemento regulador da “asa” de *D. biarmipes* ativa a expressão do gene repórter em um padrão de mancha na asa em desenvolvimento, enquanto o elemento homólogo de *D. melanogaster* não ativa o padrão de mancha de expressão de repórter. Essa diferença em atividades de elemento regulador de ação cis demonstra que as mudanças na função do elemento regulador de ação cis é subjacente a diferenças na expressão Yellow e pigmentação entre as duas espécies.

Podem ocorrer mudanças em um elemento independentemente de outros elementos, sem afetar sua função.

As mudanças evolutivas nas seqüências regulatórias de ação cis têm um papel crítico na evolução da expressão gênica. Elas evitam os efeitos pleiotrópicos das mutações nas seqüências codificantes dos genes que têm vários papéis no desenvolvimento.

A restrição nas seqüências codificantes de proteínas altamente pleiotrópicas explica a extraordinária conservação dos domínios de ligação do DNA das proteínas Hox e muitos outros fatores de transcrição em grandes períodos do tempo evolutivo. Mas, embora as funções bioquímicas das proteínas sejam restringidas, sua regulação diverge. A evolução de *Hox* e outros padrões de expressão dos genes ferramentais têm papéis importantes na evolução da forma do corpo.

Evolução regulatória em humanos

A evolução regulatória não é limitada a genes que afetam o desenvolvimento. O nível, momento ou padrão espacial da expressão de qualquer gene pode variar dentro das populações ou divergir entre as espécies. Por exemplo, como visto antes, as freqüências de alelos no locus do grupo sanguíneo Duffy variam amplamente em populações humanas (veja Quadro 19.2). Na África subsaariana, a maioria dos membros de populações indígenas têm o alelo Fy^{null} . Esse alelo resulta na ausência completa de expressão da glicoproteína Duffy nas hemácias, embora a proteína ainda esteja sendo produzida. Como e por que o antígeno Duffy está faltando nas hemácias desses indivíduos?

A explicação molecular para a falta de expressão do antígeno Duffy nas hemácias está na presença de uma mutação de ponto na região promotora do gene *Duffy* na posição -46. Essa mutação está em um sítio de ligação para um fator de transcrição específico de eritróide chamado GATA1. A mutação desse sítio abole a atividade de um acentuador do gene *Duffy* nas análises de gene repórter.

Uma explicação evolutiva sugere que a falta de expressão do antígeno Duffy entre os africanos é o resultado da seleção natural favorecendo a resistência à infecção de malária em humanos. O parasita da malária *Plasmodium vivax* é a segunda forma mais prevalente do parasita da malária nas regiões tropicais e subtropicais do mundo, mas no momento está ausente na África subsaariana. O parasita entra nas hemácias e em seus precursores ligando-se ao antígeno Duffy. A freqüência muito alta de homozigotos Fy^{null} na África impede que *P. vivax* seja comum na região. Além disso, se for suposto que *P. vivax* era comum na África no passado, então o alelo Fy^{null} teria sido favoravelmente selecionado. Como a seleção natural aumentou a freqüência do alelo, o parasita *P. vivax* teria sido cada vez menos comum porque seus hospedeiros suscetíveis seriam cada vez mais raros. Não está claro por que uma evolução similar não ocorreu na Eurásia, onde o *P. vivax* é comum e o alelo Fy^{null} é raro.

A ausência completa de antígeno Duffy nas hemácias de uma grande subpopulação implica saber se a proteína Duffy

tem alguma função, pois ela é aparentemente dispensável. Mas não ocorre que esses indivíduos também tenham a falta de expressão da proteína Duffy. A proteína é expressa nas células endoteliais do sistema vascular e células de Purkinje do cerebelo. Como na evolução da expressão Yellow nas moscas-das-frutas com manchas nas asas, a mutação regulatória do locus *Fy* permite que um aspecto da expressão gênica (nas hemácias) mude sem perturbar outros.

As modificações de seqüências codificantes e reguladoras são vias comuns da mudança evolutiva. Elas ilustram como a diversidade pode surgir sem mudar o número de genes em uma espécie. Entretanto, mudanças mutacionais em grande escala podem ocorrer e ocorrem no DNA resultando na expansão do número de genes, e essa expansão dá a matéria-prima para a inovação evolutiva.

► 19.8 A Origem de Novos Genes

Claramente, a evolução consiste em muito mais do que a simples substituição de um alelo por outro em loci com funções definidas. Surgiram novas funções que resultaram em novos modos de viver. Muitas dessas novas funções — por exemplo, o desenvolvimento do ouvido interno dos mamíferos por uma transformação dos ossos da mandíbula dos répteis — resulta de transformações contínuas de forma e não requerem genes e proteínas totalmente novos. Mas novos genes e proteínas são necessários para produzir novidades qualitativas, tais como fotossíntese e paredes celulares nas plantas, proteínas contráteis, novos tipos de células e tecidos, moléculas de oxigenação tais como hemoglobina, o sistema imunológico, ciclos de detoxificação química e enzimas digestivas. Funções metabólicas mais antigas devem ser mantidas enquanto são desenvolvidas novas, o que, por sua vez, significa que os genes antigos tinham que ser preservados enquanto novos genes com novas funções tinham que evoluir. De onde vem o DNA para novos genes?

Poliploidia

Um processo para o fornecimento de um novo DNA é a duplicação de todo o genoma por poliploidização, o que é muito mais comum em plantas do que em animais (veja o Capítulo 16). A evidência de que as poliploidias tiveram um papel importante na evolução das espécies de plantas é apresentada na Figura 19.16, que mostra a distribuição de freqüência de números de cromossomos haplóides entre espécies de plantas dicotiledôneas. Acima de um número de cromossomos de cerca de 12, os números pares são muito mais comuns do que os ímpares, uma consequência da freqüente poliploidia.

Duplicações

Um segundo modo para aumentar o DNA é pela duplicação de pequenos trechos do genoma. Tal duplicação pode ser uma consequência de má replicação do DNA. Alternativamente, um elemento de transposição pode inserir uma cópia de uma parte do genoma em outro local (veja o Capítulo 14). Após ter surgido um segmento duplicado, pode-se verificar uma

Os números pares de cromossomos são mais comuns que os ímpares

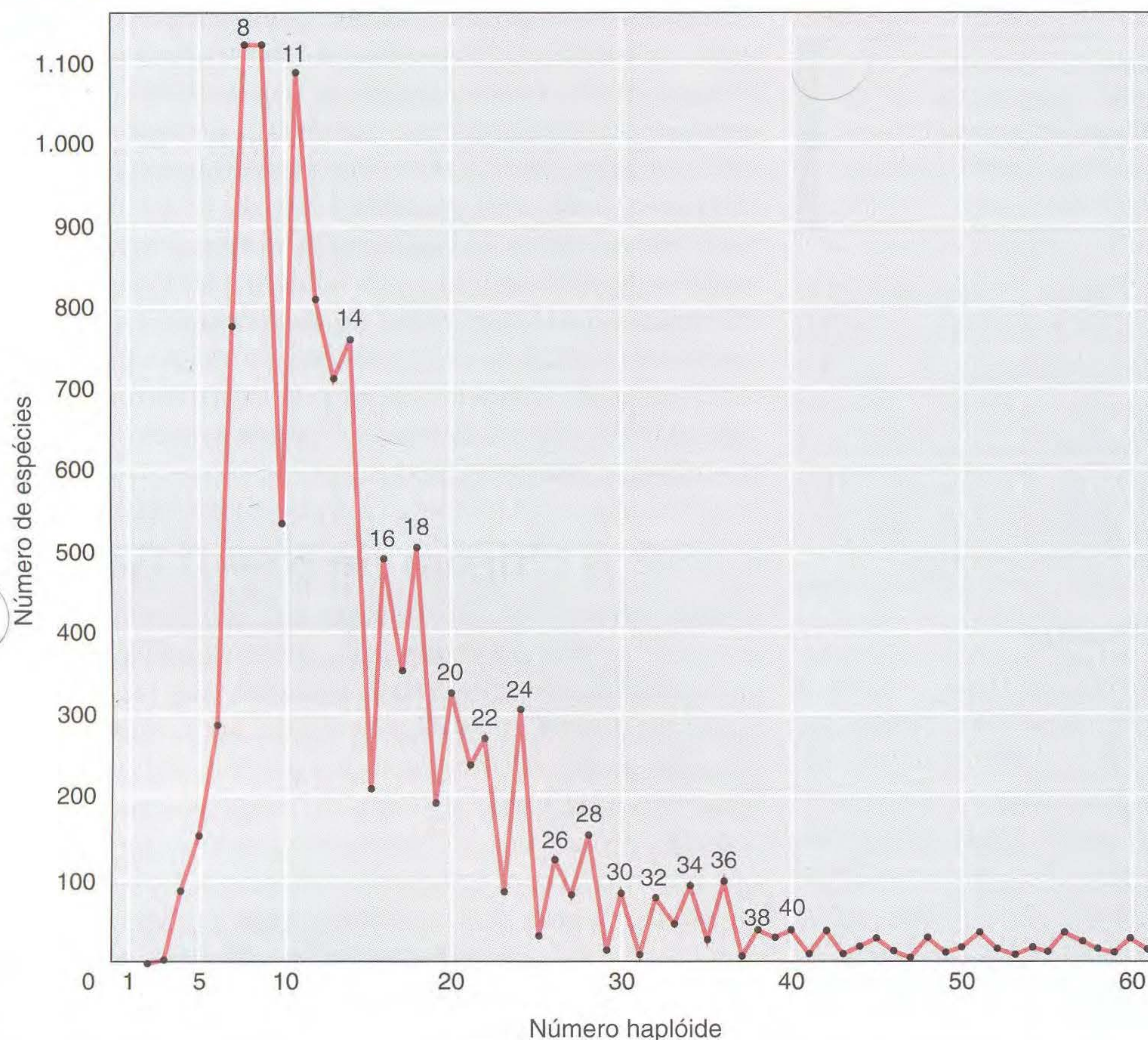


FIG. 19.16 Distribuição de frequência dos números de cromossomos haplóides em plantas dicotiledôneas. [De Verne Grant, *The Origin of Adaptations*. Columbia University Press, 1963.]

dentre três ocorrências: (1) a produção do polipeptídeo pode simplesmente aumentar; (2) a função geral da sequência original é mantida no novo DNA, mas há alguma diferenciação das sequências por mutações acumuladas, de modo que são produzidas variações na mesma proteína, permitindo uma estrutura molecular um pouco mais complexa; ou (3) o novo segmento pode divergir mais acentuadamente e ter toda uma nova função.

Um exemplo clássico do segundo caso são as duplicações gênicas e divergências subjacentes à produção da hemoglobina humana. A hemoglobina adulta é um tetrâmero que consiste em duas cadeias polipeptídicas α e duas cadeias β , cada uma com uma molécula de hemo. O gene que codifica a cadeia α está no cromossomo 16, e o gene para a cadeia β está no cromossomo 11; mas as duas cadeias são 49% idênticas em suas sequências de aminoácidos e identidade que, claramente, indicam a origem comum. Entretanto, nos fetos, até o nascimento, cerca de 80% das cadeias β são substituídas por uma cadeia γ correlata. Essas cadeias polipeptídicas β e γ são 75% idênticas. Além disso, o gene para a cadeia γ está perto do gene da cadeia β no cromossomo 11, e tem uma estrutura íntron-éxon idêntica. Essa mudança desenvolvimental na síntese de globina é parte de um grande conjunto de mudanças desenvolvimentais que são mostradas na Figura 19.17. O

embrião inicial começa com as cadeias α , γ , ϵ e ζ e, após cerca de 10 semanas, as cadeias ϵ e ζ são substituídas por α , β e γ . Perto do nascimento, β substitui γ e uma pequena quantidade de uma sexta globina, δ , é produzida.

Variações no tema hemoglobina

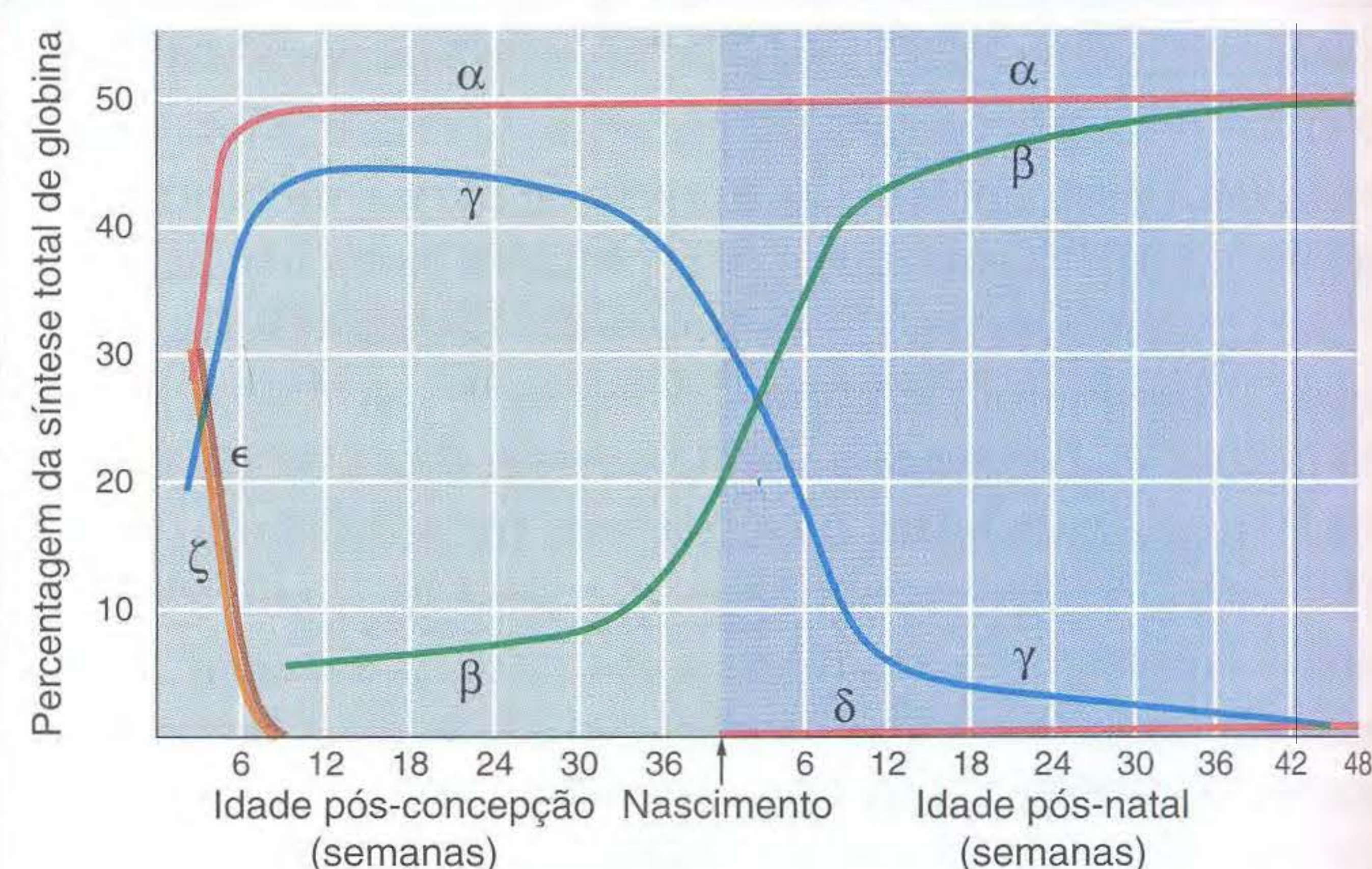


FIG. 19.17 Mudanças desenvolvimentais na síntese de globinas tipo α e tipo β que constituem a hemoglobina humana.

O Quadro 19.4 mostra a percentagem de aminoácidos identificados entre essas cadeias, e a Figura 19.18 mostra os locais cromossômicos e as estruturas íntron-éxon dos genes que os codificam. A história é marcadamente consistente. As cadeia β , δ , γ e ϵ , todas pertencem ao grupo “tipo β ”; elas têm seqüências de aminoácidos muito similares e são codificadas por genes de idêntica estrutura íntron-éxon que são todos contidos em um trecho de 60 kb de DNA no cromossomo 11. As cadeias α e ζ pertencem a um grupo “tipo α ” e são codificadas por genes contidos em uma região de 40 kb no cromossomo 16. Além disso, a Figura 19.18 mostra que, tanto no cromossomo 11 como no 16, existem pseudogenes, identificados como Ψ_α e Ψ_β . Esses pseudogenes são cópias duplicadas de genes que não adquiriram novas funções, mas acumularam mutações aleatórias que os tornaram não-funcionais. O que é marcante é que a ordem dos genes em cada cromossomo é a mesma que a ordem temporal de aparecimento das cadeias de globina no curso do desenvolvimento.

Com relação à hemoglobina, o DNA duplicado codifica uma nova proteína que desempenha uma função proxima-mente relacionada à função codificada pelo gene original. Mas o DNA duplicado pode divergir marcadamente de função. Um exemplo de tal divergência é mostrado na Figura 19.19. As aves e os mamíferos, como outros organismos eucarióticos, têm um gene codificando lisozima, uma enzi-ma protetora que degrada a parede das bactérias, como já mencionado. Esse gene foi duplicado em mamíferos para produzir uma segunda seqüência que codifica uma proteína não-enzimática completamente diferente, a α -lactalbumina, um componente nutricional do leite. A Figura 19.19 mostra que o gene duplicado tem a mesma estrutura íntron-éxon que o gene de lisozima, cujo arranjo de quatro genes e três íntrons sugere um evento anterior de múltipla duplicação na origem da lisozima.

DNA importado

As duplicações de DNA não são a única fonte de novo DNA que é a base de novas funções; ele também pode ser importa-do. Repetidamente na evolução, DNA extra tem sido impor-tado para o genoma de fontes externas por mecanismos que não a reprodução sexual normal. O DNA pode ser inserido nos cromossomos de outros locais cromossômicos e mesmo

Quadro 19.4 Percentagem de Similaridade em Seqüências de Aminoácidos entre Cadeias de Globina Humana

α	ζ	β	γ	ϵ
α	58	42	39	37
ζ		34	38	37
β			73	75
γ				80

de outras espécies. Em alguns casos, genes de organismos não-relacionados podem tornar-se incorporados a células para constituírem uma parte funcional do genoma da célula receptora.

Organelas celulares As células eucarióticas obtiveram algumas de suas organelas desse modo. Tanto os cloroplas-tos de organismos fotossintéticos quanto mitocôndrias são descendentes de procariontes que entraram nas células euca-rióticas seja como infecções ou tendo sido ingeridos. Esses procariontes tornaram-se simbiotes, transferindo grande parte de seus genomas para os núcleos de seus hospedeiros eucarióticos, mas mantendo os genes que são essenciais para as funções celulares. As mitocôndrias possuem cerca de três dúzias de genes envolvidos com a respiração celular, bem como alguns genes de tRNA; os genomas de cloroplastos, por sua vez, têm cerca de 130 genes codificando enzimas do ciclo fotossintético, bem como proteínas ribossomais e tRNA.

Evidências importantes da origem extracelular das mito-côndrias estão para ser encontradas em seu código genético. O código “universal” DNA-RNA de genes nucleares não é, de fato, universal, diferindo em alguns aspectos do das mito-côndrias. O Quadro 19.5 mostra que, para 5 das 64 trincas de RNA, as mitocôndrias diferem em seu código do genoma nuclear. Além disso, as mitocôndrias em organismos dife-rentes divergem umas das outras quanto a esses elementos codificantes, dando evidências de que as células eucarióticas devem ter sido invadidas por procariontes pelo menos cin-co vezes, cada vez por um procarionte com um sistema de codificação diferente. Para os vertebrados, vermes, insetos, o código mitocondrial é mais regular do que o código nuclear

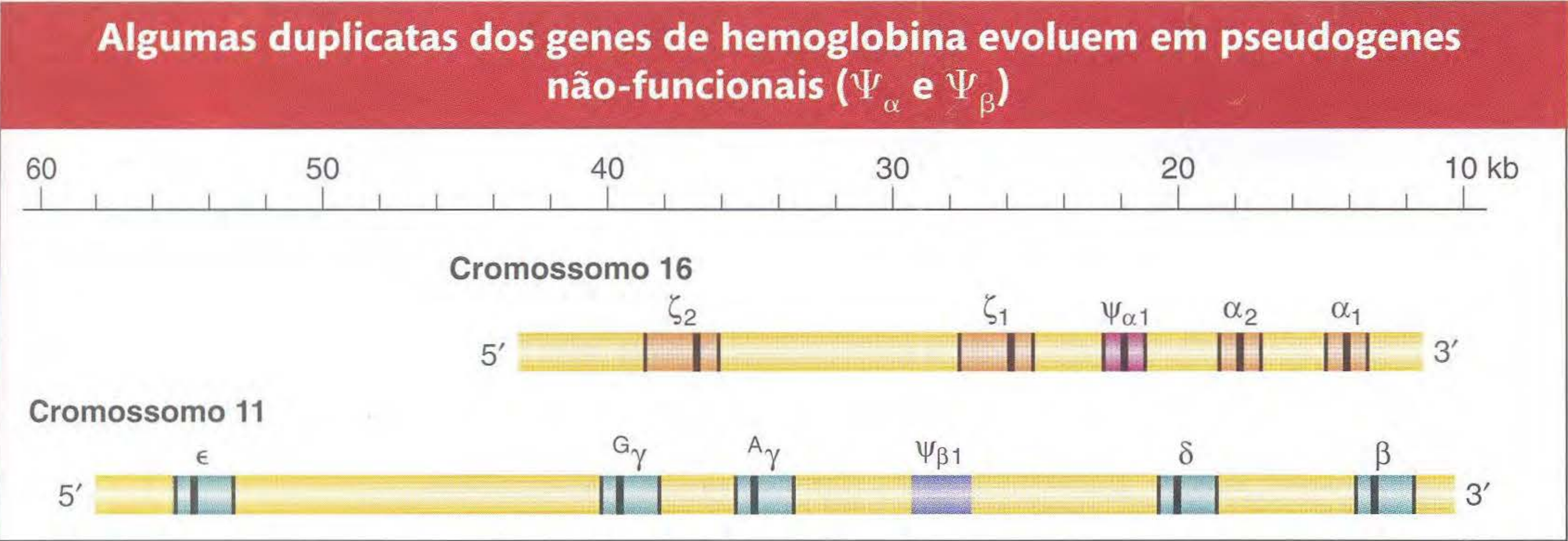


FIG. 19.18 Distribuição cromossômica de genes para a família de α -globinas no cromossomo 16 e a família de β -globinas no cromossomo 11 em humanos. A estrutura gênica é mostrada por barras pretas (éxons) e barras coloridas (íntrons).

Dois genes correlatos podem ter desenvolvido funções diferentes

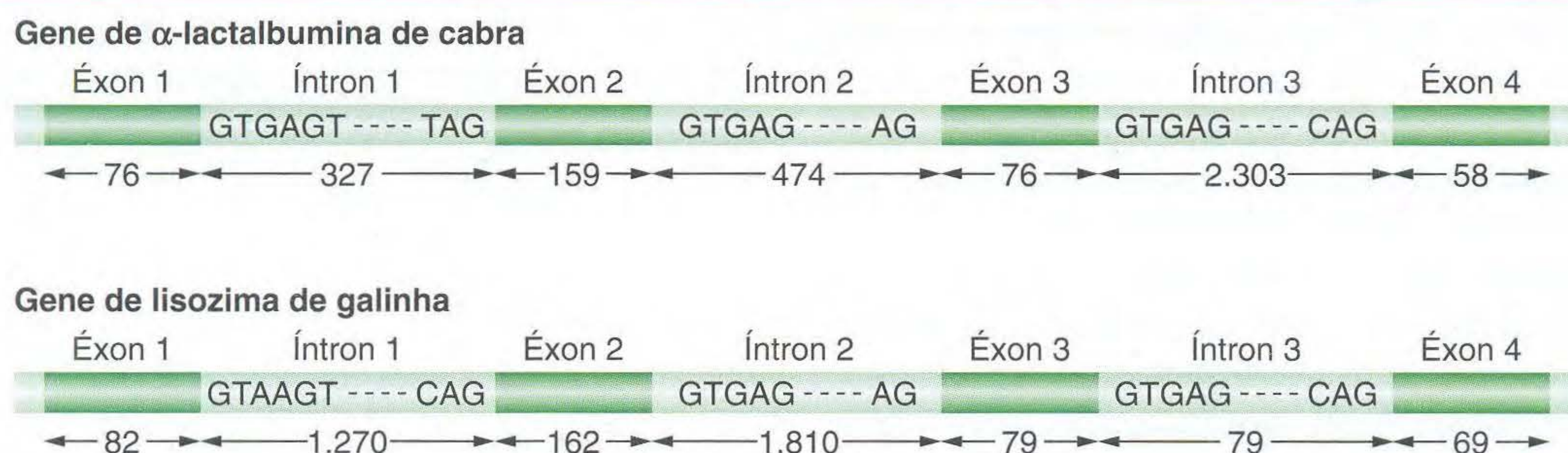


FIG. 19.19 Homologia estrutural do gene para lisozima de galinha e α -lactalbumina de mamíferos. Os éxons e íntrons são indicados por barras verde-escuras e barras verde-claras, respectivamente. As seqüências de nucleotídeos no começo e fim de cada íntron são indicadas, e os números referem-se ao tamanho de nucleotídeos de cada segmento. [De I. Kumagai, S. Takeda, and K.-I. Miura, "Functional Conversion of the Homologous Proteins α -Lactalbumin and Lysozyme by Exon Exchange," *Proc. Natl. Acad. Sci. USA* 89, 1992. 5887-5891.]

universal. No genoma nuclear, por exemplo, a isoleucina é o único aminoácido redundantemente codificado por exatamente três trincas: ATT, ATC, e ATA. A transição da terceira base de A para G produz o quarto membro desse grupo de códons, ATG, mas ele codifica metionina. Em contraste, nas mitocôndrias, esse grupo de códons contém dois códons para metionina e dois para isoleucina, separados por uma transversão.

Transferência horizontal Agora está claro que o genoma nuclear é aberto para a inserção de DNA tanto de outras partes do mesmo genoma quanto de fora. *Dentro* de um genoma, o DNA pode ser transferido pela ação de elementos de transposição (veja o Capítulo 14). Os cromossomos de uma *Drosophila*, por exemplo, contêm uma grande variedade de famílias de elementos de transposição com várias cópias de cada distribuídas pelo genoma. Até 25% do DNA de *Drosophila* podem ser originados por transposição. Não está claro qual o papel que esse DNA móvel tem na evolução funcional. Quando os elementos de transposição são introduzidos em zigotos na reprodução, tais como os elementos *P* de *Drosophila* (veja o Capítulo 14), o resultado é uma proliferação explosiva dos elementos no genoma do receptor. Quando um

elemento móvel é inserido em um gene, o efeito no organismo geralmente é marcante e deletério, mas esse efeito pode ser um artefato dos métodos usados para detectar a presença de tais elementos. Os resultados de experimentos de seleção em laboratório sobre características quantitativas mostraram que a transposição pode agir como uma fonte adicional de variação selecionável. Há também a possibilidade de que os genes sejam transferidos do genoma nuclear de uma espécie para o genoma nuclear de outra por retrovírus (veja o Capítulo 14). Os retrovírus podem ser levados entre espécies muito distantemente relacionadas por vetores comuns de doença, tais como insetos, ou por infecções bacterianas. Assim, qualquer material genético exógeno levado por um retrovírus pode ser uma poderosa fonte de novas funções.

► 19.9 Evidência Genética de Ancestralidade Comum na Evolução

Quando pensamos em evolução, pensamos em mudança. As espécies que vivem em determinada época são diferentes

Quadro 19.5 Comparação do Código de DNA Nuclear Universal com Vários Códigos Mitocondriais para Cinco Trincas nos quais Eles Diferem

	Códigos de trincas				
	TGA	ATA	AGA	AGG	AAA
Nuclear	Fim	Ile	Arg	Arg	Lis
Mitocondrial					
Mamíferos	Trp	Met	Fim	Fim	Lis
Aves	Trp	Met	Fim	Fim	Lis
Anfíbios	Trp	Met	Fim	Fim	Lis
Equinodermas	Trp	Ile	Ser	Ser	Asn
Insetos	Trp	Met	Ser	Fim	Lis
Nematódeos	Trp	Met	Ser	Ser	Lis
Platelmintos	Trp	Met	Ser	Ser	Asn
Cnidários	Trp	Ile	Arg	Arg	Lis

de seus ancestrais, tendo mudado de forma e função pelos mecanismos revistos até agora na discussão da genética do processo evolutivo. Mas existe uma segunda característica da diversidade da vida, aquela que Darwin usou como argumento importante da realidade da evolução. Não só os organismos atuais descendem de organismos anteriores, diferentes, mas, se voltarmos no tempo, os organismos que atualmente são muito diferentes descendem de uma única forma ancestral. De fato, se voltarmos bastante no tempo para a origem da vida, todos os organismos na Terra são descendentes de um único ancestral comum. Assim, esperamos observar que, aparentemente, espécies diferentes têm similaridades subjacentes, atributos de seu ancestral comum que foram conservados ao longo do tempo evolutivo a despeito de todas as mudanças que ocorreram.

Antes de haver disponibilidade da bioquímica e genética moderna, a principal evidência de similaridade subjacente de estruturas aparentemente diferentes em espécies diferentes foi obtida de observações anatômicas de formas embrionárias e adultas. Assim, estruturas ósseas similares de ossos das asas de morcegos e membros anteriores de mamíferos corredores tornaram evidente que essas estruturas eram evolutivamente derivadas de um mamífero ancestral comum. Além disso, a anatomia das asas dos pássaros indica a ancestralidade comum dos mamíferos e aves (Figura 19.20).

Como indicado na discussão das ferramentas genéticas para o desenvolvimento no Capítulo 12, a análise genética do desenvolvimento deu poderosas demonstrações da ancestralidade comum de animais tão diferentes quanto insetos e mamíferos. A explicação mais simples para os conjuntos similares de genes entre animais é que ferramentas genéticas (*toolkit*) completas estavam presentes em um ancestral comum há cerca de 600 milhões de anos. Muitos exemplos foram descobertos de genes fortemente conservados, e mesmo vias inteiras que têm função similar. Tal conservação evolutiva e funcional parece ser a norma e não a exceção. O que tornou a genética do desenvolvimento um campo extraordinariamente excitante de pesquisas biológicas é a demonstração, por meio de análise genética, de que as vias básicas de desenvolvimento e sua base genética foram conservadas por centenas de milhões de anos de evolução.

Estratégias desenvolvimentais em animais são muito antigas e altamente conservadas. Em essência, um mamífero, um verme e uma mosca são colocados juntos com os mesmos blocos estruturais genéticos básicos e modos de regulação. *Plus ça change, plus c'est le même chose!* (Quanto mais muda, mais é a mesma coisa!)

Em um segundo nível, mais profundo, podemos observar a origem evolutiva comum de organismos na estrutura de suas proteínas e seus genomas. A vantagem da observação direta das seqüências da proteína e do DNA é que não temos que depender de observar a similaridade de função entre as proteínas ou estruturas anatômicas que resultam da posse de determinados genes. Já vimos que substituir um único aminoácido pode mudar a função de uma proteína de uma

A anatomia indica ancestrais comuns

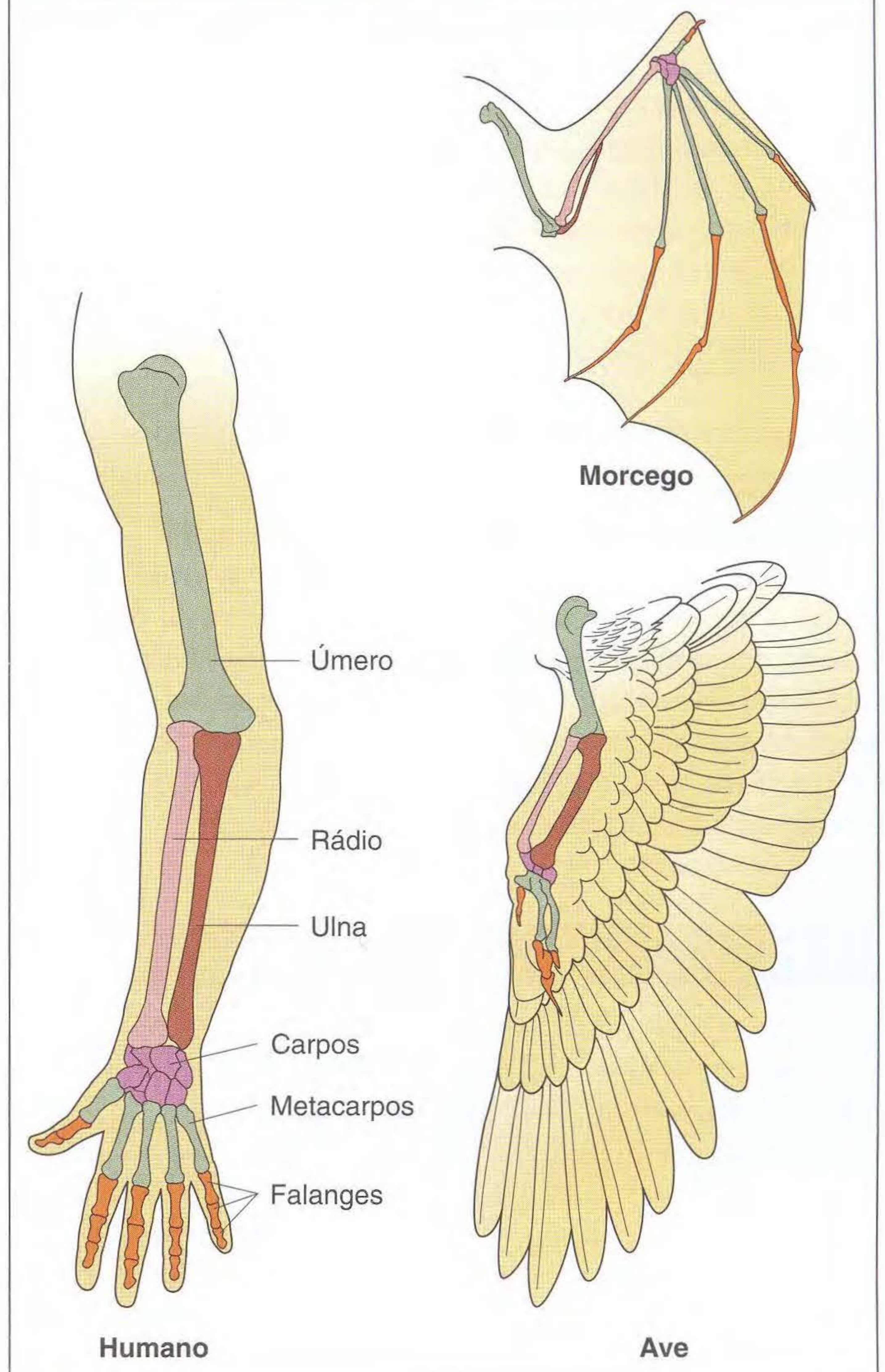


FIG. 19.20 As estruturas ósseas de uma asa de morcego, uma asa de pássaro e um braço e mão humanos. Essas estruturas ósseas mostram a similaridade anatômica subjacente entre elas e o modo pelo qual ossos diferentes foram relativamente aumentados ou diminuídos para produzir essas estruturas diferentes. [De W.T. Keeton and J.L. Gould, *Biological Science*. Norton, 1986.]

esterase para uma fosfatase ácida. Mas, a despeito dessa mudança de função, não temos dificuldade em determinar que as duas enzimas são produzidas pela leitura dos genes que são praticamente idênticos, um dos quais foi derivado de uma única etapa mutacional do outro, como a resistência a inseticidas evoluiu por seleção natural.

Ao longo do tempo evolutivo, os genes que descenderam de um ancestral comum irão divergir em seqüência de DNA e na sua posição física no genoma, como resultado de mutações e rearranjos cromossômicos. Se passar um tempo suficiente e não houver forças contrárias de seleção natural, essa divergência finalmente resultará na perda de qualquer similaridade observável nos genes ou proteínas entre espécies diferentes, mesmo que descendam de um ancestral comum.

De fato, nem mesmo o tempo transcorrido desde os ancestrais comuns dos atuais vertebrados e invertebrados apagou a similaridade das seqüências de DNA e aminoácidos entre *Drosophila* e camundongos. Não apenas as taxas de mutação não são altas o suficiente para causar a perda completa de similaridade mesmo em centenas de milhões de anos, como também a maioria das novas mutações não é preservada, porque causam uma perda deletéria ou mudança de função de uma proteína ou no controle do tempo e lugar de produção de proteína. Assim, a quantidade de divergência que foi preservada na evolução tem sido limitada.

Comparando os proteomas entre espécies distantes

Um grande esforço da genética molecular é dirigido para determinar a seqüência completa de DNA de uma variedade de espécies diferentes. Na época em que este parágrafo foi escrito, centenas de genomas tinham sido seqüenciados, incluindo os de muitas bactérias, fungos, plantas e animais. Enquanto você está lendo estas linhas, muito mais genomas de mais espécies terão sido seqüenciados. A disponibilidade de tais dados possibilita reconstruir a evolução de genomas

de muitas espécies diversas de seus ancestrais comuns. Além disso, hoje é possível deduzir as similaridades e diferenças nos proteomas dessas espécies comparando as seqüências de genes de várias espécies com as seqüências de genes que codificam seqüências de aminoácidos de proteínas com função conhecida.

Em nosso estado atual de conhecimentos, podemos sugerir funções para cerca de metade das proteínas no proteoma de cada um dos eucariontes cujos genomas foram seqüenciados, usando a similaridade de suas seqüências com proteínas de função conhecida. A Figura 19.21 mostra a distribuição dessa metade de cada proteoma nas categorias funcionais gerais. Marcantemente, o grupo de proteínas envolvido na defesa e imunidade se expandiu muito nos humanos em comparação com o de outras espécies. Para outras categorias funcionais, embora existam números maiores de proteínas na linhagem humana, não existe um caso em que as diferenças entre humanos e todos os outros eucariontes sejam tão pronunciadas. Como discutido nos Capítulos 10 e 11, a expressão gênica é geralmente controlada pela regulação ou transcrição por proteínas chamadas de fatores de transcrição. Talvez como uma manifestação de muitos tipos celulares que

Distribuição de proteínas por função

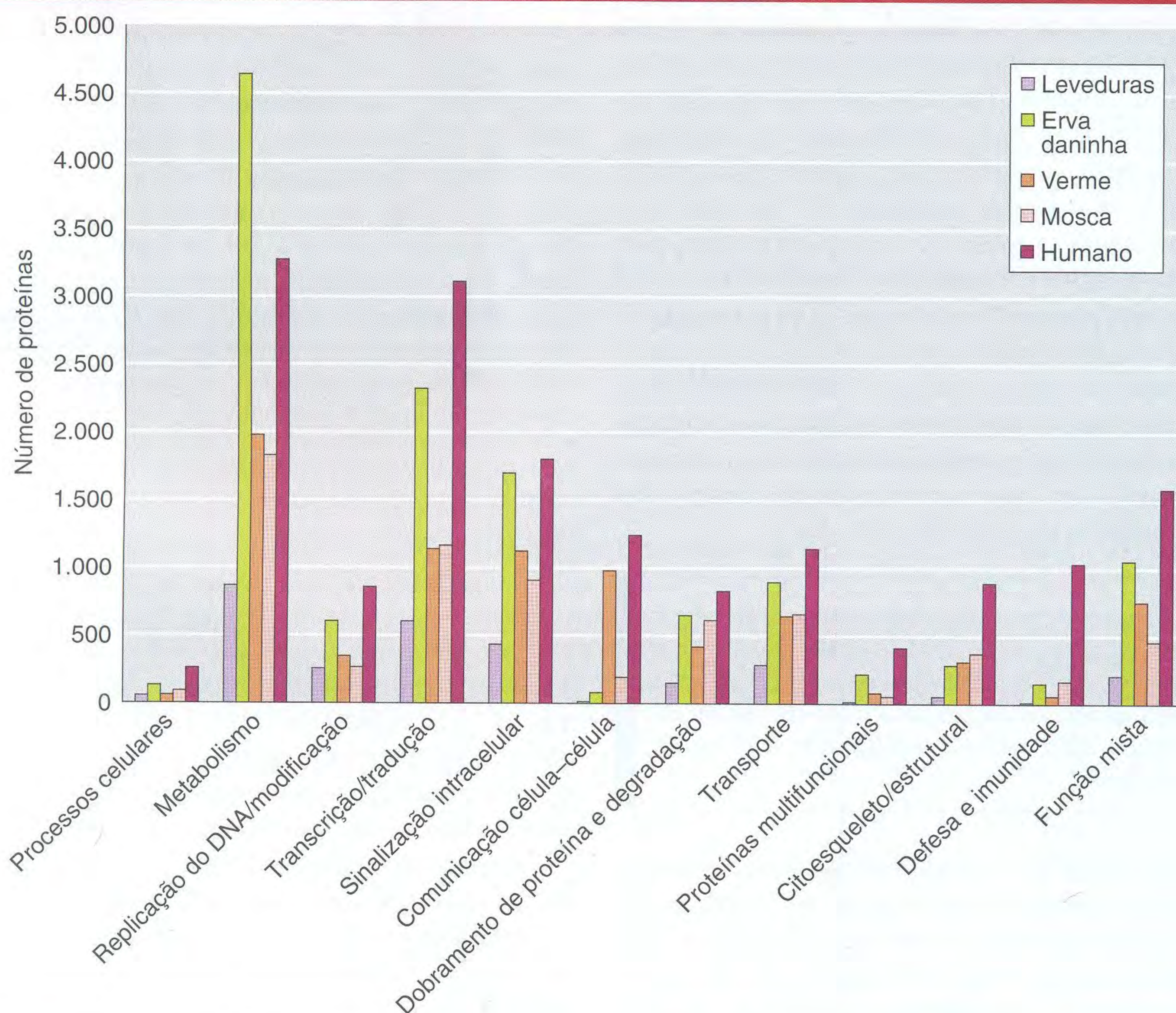


FIG. 19.21 A distribuição de proteínas eucarióticas de acordo com amplas categorias de funções biológicas. [Reimpressa com permissão de Nature 409 (15 February 2001), 902, "Initial Sequencing and Analysis of the Human Genome," The International Human Genome Sequencing Consortium. Copyright 2001 Macmillan Magazines Ltd.]

se diferenciam em humanos, o tamanho e distribuição das famílias de fatores específicos de transcrição em humanos excedem muito os números de outros eucariontes seqüenciados, com a exceção da mostardeira (*Arabidopsis thaliana*; veja Figura 19.21).

A distribuição de proteínas descritas no parágrafo anterior é uma descrição apenas de metade de cada proteoma. E quanto à outra metade? Ela pode ser dividida em dois componentes. Um componente, compreendendo cerca de 30% de cada proteoma, consiste em proteínas que têm correlatas entre genomas diferentes, mas nenhuma teve uma função atribuída a ela. O outro componente, constituindo os cerca de 20% restantes de cada proteoma, consiste em proteínas que não são relacionadas por seqüência de aminoácidos a nenhuma proteína conhecida em outro ramo da árvore evolutiva eucariótica. Podemos imaginar duas explicações possíveis para esses novos polipeptídeos. Uma possibilidade é que alguns desses polipeptídeos tenham primeiro evoluído após as espécies seqüenciadas, tendo um ancestral comum, afastarem-se umas das outras. Como nenhuma dessas espécies é evolutivamente mais próxima que algumas centenas de milhões de anos, talvez não seja surpreendente encontrar essa frequência de proteínas recém-evoluídas. A outra possibilidade é que algumas dessas proteínas estejam evoluindo muito rapidamente, e, assim, sua ancestralidade tenha sido essencialmente apagada pela sobreposição de novas mutações que se acumularam. É quase certo que ambas as possibilidades estejam corretas para um subgrupo desses novos polipeptídeos.

Finalmente, podemos perguntar: De onde vieram os genes codificantes de proteínas no genoma humano? A Figura 19.22 mostra a distribuição de genes humanos em outras

espécies. Cerca de um quinto dos genes humanos conhecidos foi encontrado apenas em vertebrados. Outro quinto parece ser ubíquo em eucariontes e procariontes. Cerca de um terço é encontrado nos eucariontes, mas não em bactérias. Mas, quando o genoma humano é comparado com o de outros mamíferos, vemos que a grande maioria dos genes são comuns a todos eles, indicando que o último ancestral comum de todos os mamíferos tinha a maioria dos genes que os humanos têm hoje em dia.

Comparação dos genomas entre vizinhos próximos: Genômica comparativa humano-camundongo

Os genomas são tidos como tendo evoluído em parte por um processo de rearranjos cromossômicos — isto é, quebrando e reunindo os arcabouços de moléculas bifilamentares de DNA, produzindo assim novas ordens de genes e novos cromossomos. (Veja o Capítulo 16 para uma discussão de rearranjos cromossômicos.) A extensão com a qual os rearranjos cromossômicos acumularam-se durante a evolução pode ser avaliada procurando ordens comuns de genes entre espécies divergentes. Como exemplo desse enfoque, compararemos os genomas de humanos e camundongos, duas espécies que divergiram de um ancestral comum há cerca de 50 milhões de anos. O genoma do camundongo já foi suficientemente seqüenciado para que as ordens relativas dos genes possam ser determinadas. Grandes blocos de ordem de genes conservados são facilmente reconhecidos. Por meio de comparações sistemáticas desse tipo, podemos fazer **mapas de sintenia**, que mostram a origem cromossômica de uma espécie essencialmente colorida sobre o cariótipo da outra. A Figura 19.23 mostra uma representação em cores do genoma sintênico camundongo-humano. Nessa ilustração, 21 cores diferentes representam os cromossomos sexuais X e Y de camundongo e os 19 autossomos. Por exemplo, a maioria dos cromossomos 14 de camundongo podem ser encontrados em três blocos do cromossomo humano 13, mas, além disso, pequenos segmentos podem ser encontrados nos cromossomos humanos 3, 8, 10 e 14. Distribuições similares bloco a bloco no genoma humano são observadas para cada um dos outros cromossomos de camundongo. Assim, podemos concluir que muitos rearranjos cromossômicos ocorreram entre os seres humanos e o camundongo, mas não o suficiente para ter embaralhado os dois genomas um em relação ao outro.

A genômica comparativa é uma fonte de *insight* das mudanças em níveis gênico e cromossômico que ocorreram no processo de evolução.

Distribuição de genes humanos em outras espécies

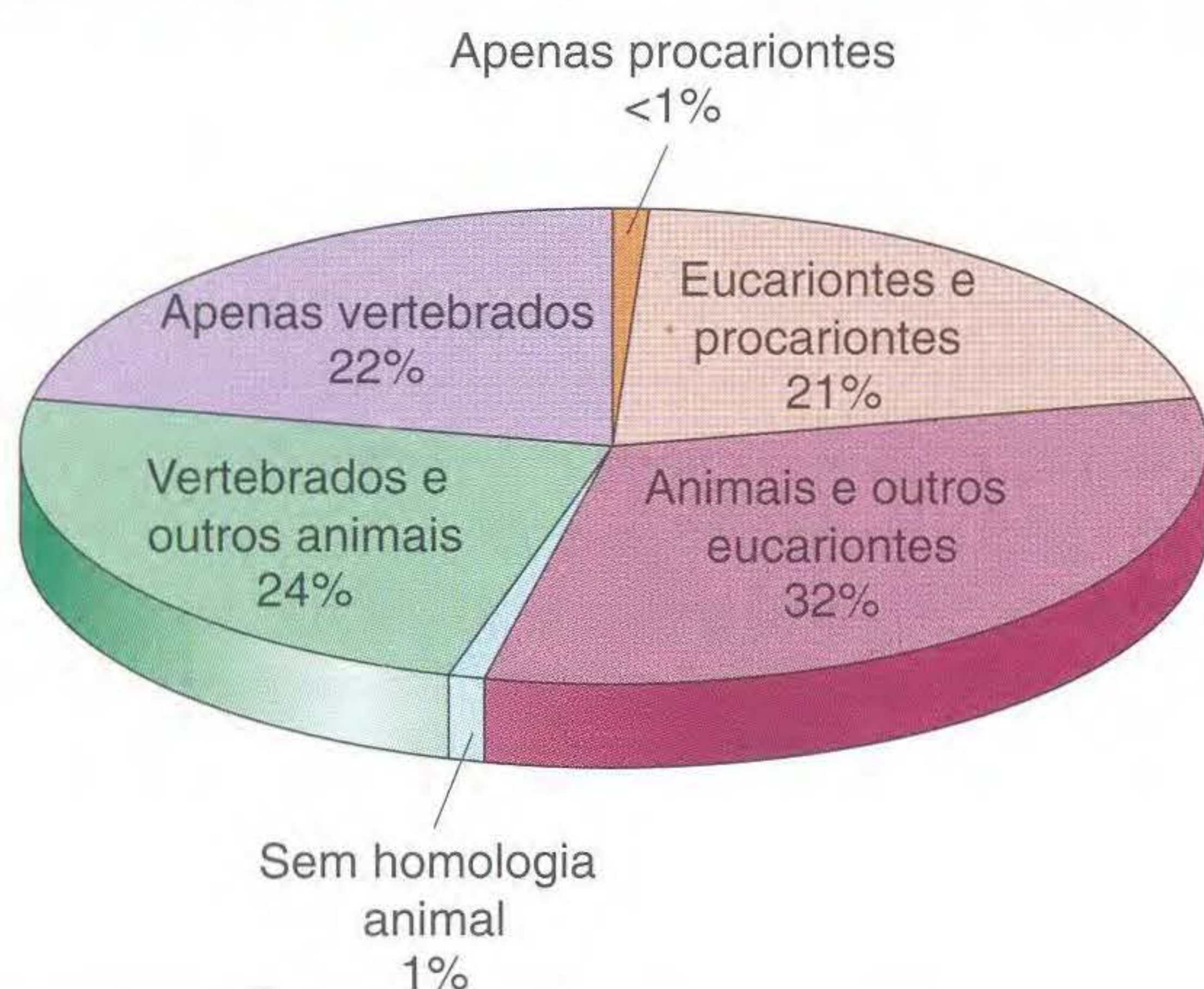


FIG. 19.22 A distribuição de proteínas humanas de acordo com a identificação de proteínas significativamente relacionadas em outras espécies. Note que cerca de um quinto das proteínas humanas foi identificado apenas dentro da linhagem de vertebrados, enquanto, no outro extremo, um quinto foi identificado em todos os principais ramos da árvore evolutiva. [Reimpresso com permissão de *Nature* 409 (15 February 2001), 902, "Initial Sequencing and Analysis of the Human Genome," The International Human Genome Sequencing Consortium. Copyright 2001 Macmillan Magazines Ltd.]

► 19.10 O Processo de Especiação

Quando examinamos o mundo vivo, vemos que organismos individuais geralmente são agrupados em coleções que se assemelham umas às outras mais ou menos proximamente, e são claramente distintas de outros grupos. Um exame íntimo de uma prole de *Drosophila* mostrará diferenças no número

Um mapa de sintenia do genoma humano

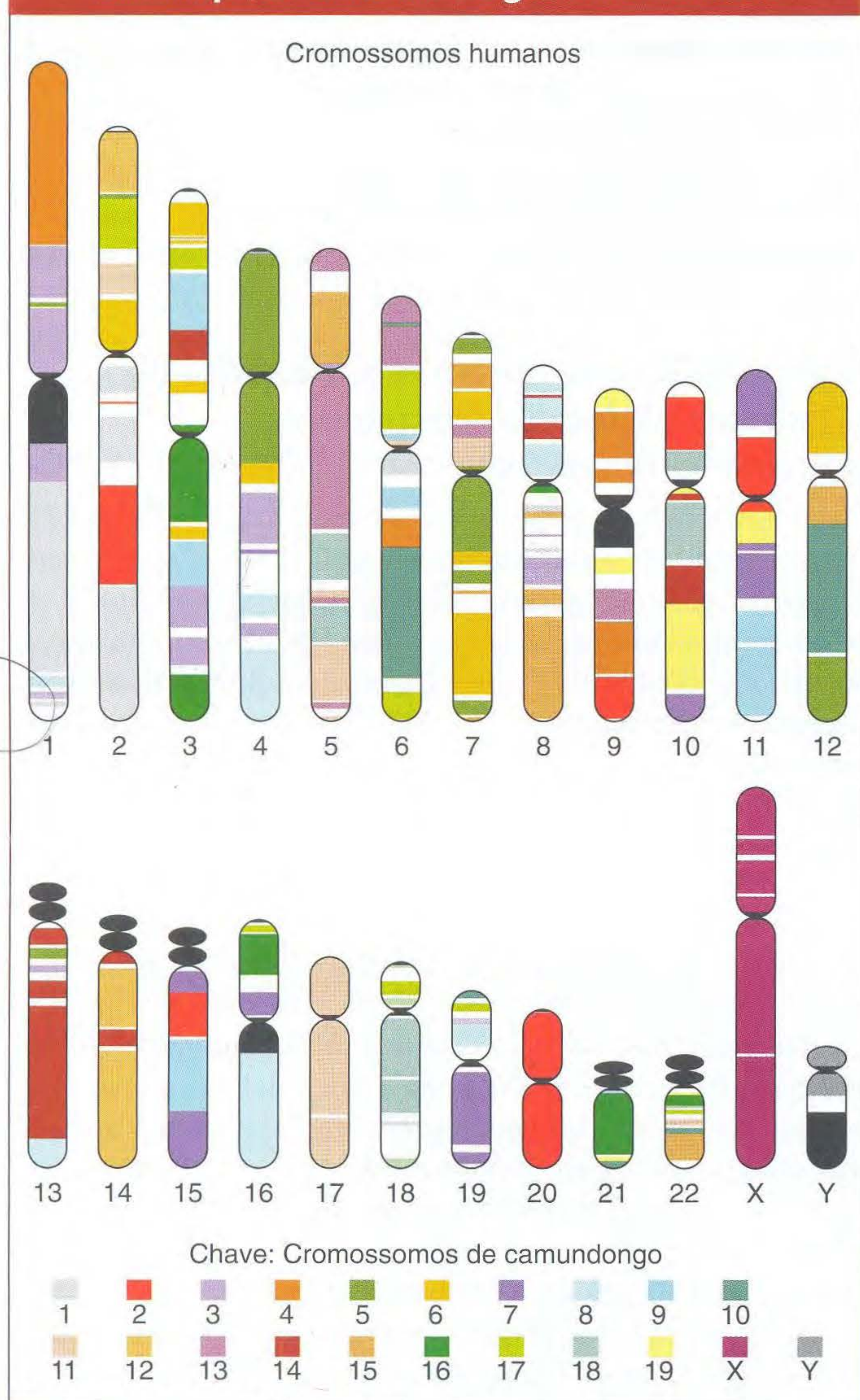


FIG. 19.23 Este mapa de sintenia usa um código de cores para mostrar correspondências regionais de cada bloco do genoma humano com as seções correspondentes do genoma de camundongo. Cada cor representa um cromossomo diferente de camundongo, como indicado na chave. [Reimpresso com permissão de *Nature* 409 (15 February 2001), 910, "Initial Sequencing and Analysis of the Human Genome," The International Human Genome Sequencing Consortium. Copyright 2001 Macmillan Magazines Ltd.]

de cerdas, tamanho do olho e detalhes de padrões de cor de mosca a mosca, mas um entomologista não terá nenhuma dificuldade em distinguir *Drosophila melanogaster* de, digamos, *Drosophila pseudoobscura*. Nunca vemos uma mosca que seja intermediária a esses dois tipos. Claramente, pelo menos na natureza, não há entrecruzamento efetivo de duas dessas formas. Um grupo de organismos que troca genes dentro do grupo mas não pode fazer isso com outros grupos é o que caracteriza uma **espécie**. Dentro das espécies podem existir populações locais que também são facilmente distintas umas das outras por algumas características fenotípicas, mas

também os genes podem ser facilmente trocados entre eles. Por exemplo, ninguém tem dificuldade em distinguir um senegalês "típico" de um sueco "típico", mas tais pessoas são capazes de se reproduzir umas com as outras e gerar prole. De fato, têm havido muitas dessas reproduções na América do Norte nos últimos 300 anos, criando um enorme número de pessoas de vários graus intermediários entre esses tipos geográficos locais. Eles não são espécies separadas. Em geral, existem algumas diferenças na frequência de vários genes em populações geográficas diferentes de qualquer espécie; assim, marcar determinada população como uma raça distinta é arbitrário e, como consequência, o conceito de raça não é mais muito usado em biologia.

Uma espécie é um grupo de organismos que podem trocar genes entre si mas que são geneticamente incapazes de trocar genes na natureza com indivíduos de outros grupos. Uma raça geográfica é uma população local fenotipicamente distinguível dentro de uma espécie e que é capaz de trocar genes com outras raças dentro dessa espécie. Como quase todas as populações geográficas são diferentes de outras nas frequências de alguns genes, raça é um conceito que não faz distinção biológica clara.

Todas as espécies hoje existentes estão relacionadas umas às outras, tendo tido um ancestral comum em alguma época do passado evolutivo. Isso significa que cada uma dessas espécies separou-se de uma espécie previamente existente e tornou-se geneticamente distinta e geneticamente isolada de sua linhagem ancestral. Em circunstâncias extraordinárias, uma única mutação pode ser suficiente para encontrar tal grupo geneticamente isolado, mas o portador dessa mutação precisaria ser capaz de autofecundação ou reprodução vegetativa. Além disso, essa mutação teria que causar uma incompatibilidade completa entre seu portador e a espécie original, e permitir que a nova linhagem competisse bem-sucedidamente com o grupo previamente estabelecido. Embora não sejam impossíveis, tais eventos devem ser raros.

Mais comumente, novas espécies formam-se como um resultado de isolamento geográfico. Já vimos como populações que são geograficamente separadas irão divergir umas das outras geneticamente como consequência de mutações únicas, seleção e deriva genética. A migração entre as populações evitará que elas venham a divergir muito. Até mesmo um único migrante por geração é suficiente para evitar que as populações fixem alelos alternativos apenas por deriva genética, e mesmo a seleção para diferentes picos adaptativos não terá sucesso em causar divergência completa, a menos que seja extremamente forte. Como consequência, as populações que divergem o suficiente para se tornarem novas espécies, reprodutivamente isoladas, devem primeiro ficar totalmente isoladas umas das outras por alguma barreira mecânica. Esse isolamento quase sempre requer alguma separação espacial, e a separação deve ser grande o suficiente, ou as barreiras naturais à passagem de migrantes devem ser fortes o suficiente para impedir qualquer migração efetiva. Tais populações isoladas espacialmente são chamadas de

alopátricas. A barreira de isolamento deve ser, por exemplo, o prolongamento de uma geleira continental durante épocas glaciais que força a separação de populações antes distribuídas continuamente, ou a separação de continentes que se tornaram separados por água, ou a infrequente colonização de ilhas que estão distantes da costa. O ponto crítico é que essas barreiras devem fazer com que uma posterior migração entre populações separadas seja um evento raro. Assim, as populações agora são geneticamente independentes e continuarão a divergir por mutação, seleção e deriva genética. Ocasionalmente, a diferenciação genética entre as populações torna-se tão grande que a formação de híbridos entre elas seria fisiológica, desenvolvimental ou comportamentalmente impossível mesmo se a separação geográfica fosse abolida. Essas populações *biologicamente* isoladas agora são novas espécies, formadas pelo processo de **especiação alopátrica**.

A especiação alopátrica ocorre por um isolamento inicial geográfico e mecânico de populações que impedem o fluxo gênico entre elas, seguido de divergência genética das populações isoladas suficiente para tornar biologicamente impossível que elas venham a trocar genes no futuro.

Existem dois principais mecanismos de isolamento biológico: mecanismos de isolamento pré-zigótico mecânico e mecanismos pós-zigóticos. O **isolamento pré-zigótico** ocorre quando há uma falha em formar zigotos. A causa dessa falha pode ser que espécies diferentes se cruzem em estações diferentes do ano ou em habitats diferentes. Pode ocorrer também que as espécies não sejam sexualmente atraentes uma para a outra, ou suas genitálias não se ajustem, ou os gametas masculinos sejam fisiologicamente incompatíveis com os da fêmea.

Os exemplos de mecanismos de isolamento pré-zigótico são bem conhecidos em plantas e animais. As duas espécies de pinheiros que crescem na península de Monterey, *Pinus radiata* e *P. muricata*, espalham seus polens em fevereiro e abril, respectivamente, e, assim, não trocam genes. Os sinais luminosos que são emitidos pelos vagalumes machos e atraem as fêmeas diferem em intensidade e tempo entre as espécies. Na mosca tsé-tsé, *Glossina*, as incompatibilidades mecânicas causam grave dano, e mesmo a morte, se machos de uma espécie cruzam com fêmeas de outra. O pólen de

espécies diferentes de *Nicotiana*, o gênero ao qual pertence o tabaco, ou não germina ou não pode crescer no estilo de outras espécies. O **isolamento pós-zigótico** resulta da falha de zigotos fertilizados em contribuir com gametas para futuras gerações. Os híbridos podem falhar em desenvolver-se ou ter uma probabilidade menor de sobrevivência que a das espécies parentais, ou os híbridos podem ser parcial ou completamente estéreis. O isolamento pós-zigótico é mais comum em animais do que em plantas, aparentemente porque o desenvolvimento de muitas plantas é muito mais tolerante a incompatibilidades genéticas e variações cromossômicas. Quando os ovócitos da rã-leopardo, *Rana pipiens*, são fertilizados pelos espermatozoides da rã-madeira, *R. sylvatica*, os embriões não se desenvolvem. Cavalos e jumentos podem ser facilmente cruzados para produzir as mulas, mas, como é bem conhecido, esses híbridos são estéreis.

Genética do isolamento de espécies

Geralmente, não é possível fazer nenhuma análise genética dos mecanismos de isolamento entre duas espécies pelo simples motivo que, por definição, eles não podem ser cruzados uns com os outros. É possível, entretanto, fazer uso de espécies proximamente relacionadas nas quais o mecanismo de isolamento não produziu uma esterilidade híbrida completa e degradação dos híbridos (*hybrid breakdown*). Essas espécies podem ser cruzadas e a prole segregante de híbridos F_2 ou gerações de retrocruzamento podem ser analisadas usando marcadores genéticos e a técnica de localizar loci de características quantitativas (QTL) discutidos no Capítulo 18. Quando tais experimentos marcadores foram feitos em outras espécies, principalmente no gênero *Drosophila*, as conclusões gerais são que diferenças gênicas responsáveis pela inviabilidade híbrida estão em todos os cromossomos mais ou menos igualmente e que, para a esterilidade híbrida, existe algum efeito adicional do cromossomo X. Para o comportamento de isolamento sexual, os resultados são variáveis. Em *Drosophila*, todos os cromossomos estão envolvidos, mas, em Lepidoptera, os genes estão muito mais localizados, aparentemente devido ao envolvimento de ferormônios específicos cujo odor é importante no reconhecimento da espécie. O cromossomo sexual tem um efeito muito forte nas borboletas; na broca do milho europeu, por exemplo, apenas três loci, um dos quais está no cromossomo sexual, contribuem para todo o isolamento entre os tipos de ferormônios dentro da espécie.

► Resumo

A teoria darwiniana da evolução explica as mudanças que ocorrem em populações de organismos como resultado de mudanças nas frequências relativas de variantes diferentes na população. Se não houver variação dentro de uma espécie para alguma característica, não pode haver evolução. Além disso, essa variação deve ser influenciada por diferenças genéticas. Se as diferenças não forem herdáveis, elas não podem evoluir, pois a reprodução diferencial de variantes

diferentes não pode ser levada por gerações de linhagens. Assim, todas as reconstruções evolutivas hipotéticas dependem criticamente de se as características em questão são, de fato, herdáveis. Os processos que dão origem à variação dentro da população são causalmente independentes dos processos que são responsáveis pela reprodução diferencial dos vários tipos. É essa independência que é significativa quando se diz que as mutações são “aleatórias”. O processo

de mutação fornece uma variação não-dirigida, enquanto o processo de seleção natural seleciona essa variação, aumentando a frequência dessas variantes que, por acaso, são mais capazes de sobreviver e se reproduzir. Muitas são colhidas, mas poucas são selecionadas.

A divergência evolutiva das populações no espaço e no tempo não é apenas uma consequência da seleção natural. A seleção natural não é um processo globalmente otimizado que encontra os “melhores” organismos para determinado ambiente. Em vez disso, ela encontra um conjunto de “boas” soluções alternativas para problemas adaptativos, e o resultado particular da evolução seletiva em determinado caso é sujeito a eventos históricos casuais. Fatores aleatórios, tais como deriva genética e ocorrência casual ou perda de novas mutações, podem resultar em resultados radicalmente diferentes de um processo evolutivo mesmo quando a força da seleção natural é a mesma. A metáfora geralmente empregada é que há um “panorama adaptativo” de combinações genéticas, e que a seleção natural leva a população a um “pico” nesse panorama, mas apenas para um dos vários picos locais alternativos.

Nem toda a evolução é impulsionada pelas forças de seleção natural. Se a diferença seletiva entre duas variantes genéticas for suficientemente pequena, menos que o recíproco do tamanho da população, pode haver uma substituição de um alelo por outro puramente por deriva genética. Uma grande parte da evolução molecular parece ser a substituição de uma sequência de proteína por outra de função equivalente. A evidência dessa evolução neutra é que o número de diferenças de aminoácidos entre duas espécies diferentes em alguma molécula — por exemplo, a hemoglobina — é diretamente proporcional ao número de gerações desde sua divergência de um ancestral comum no passado evolutivo. Tal “relógio molecular” com uma taxa constante de mudança não seria esperado se a seleção das diferenças fosse dependente de mudanças particulares no ambiente. Além disso, seria esperado que o relógio andasse mais rápido para proteínas tais como fibrinopeptídeos, nas quais a composição de aminoácidos não é crítica para a função, e essa diferença na velocidade do relógio é de fato observada. Assim, não podemos supor sem evidência que as mudanças evolutivas são o resultado de seleção natural adaptativa.

Tanta evolução de sequências é efetivamente neutra que não há uma relação simples entre a quantidade de mudanças na sequência de DNA de um gene e a quantidade de mudança, se alguma, na função da proteína codificada. Algumas funções de proteínas podem mudar por uma única substituição de aminoácido, enquanto outras precisam de um conjunto de substituições. Uma restrição importante na evolução das sequências codificantes é o efeito pleiotrópico das mutações. Se uma proteína tem múltiplas funções em tecidos diferentes, as mutações nas sequências codificantes

podem afetar todas as funções e ter consequências negativas para a adaptabilidade. Os potenciais efeitos pleiotrópicos das mutações codificantes podem ser evitados por mutações em elementos regulatórios não-codificantes que podem mudar seletivamente a expressão de um gene apenas em um tecido ou parte do corpo e não em outras. A evolução da sequência regulatória de ação cis é central para a evolução de características morfológicas e a expressão dos genes ferramentais que controlam o desenvolvimento.

Novas funções geralmente surgem pela evolução de genes duplicados. Esse novo DNA pode surgir pela duplicação do genoma inteiro (poliploidia) seguida de uma lenta divergência evolutiva do conjunto cromossômico extra, o que tem sido uma frequente ocorrência em plantas. Uma alternativa é a duplicação de genes isolados seguida de seleção da diferenciação. Uma outra fonte de DNA, recentemente descoberta, é a entrada no genoma de DNA de organismos totalmente não-relacionados por infecção seguida de integração do DNA exógeno no genoma nuclear, ou pela formação de organelas celulares extranucleares com seus próprios genomas. As mitocôndrias e os cloroplastos em organismos superiores surgiram por essa via.

A grande diversidade de formas de vida diferentes que existem é uma consequência das histórias evolutivas independentes de populações separadas. Para populações diferentes divergirem uma da outra, elas não devem trocar genes; assim, a evolução independente de grandes números de espécies diferentes requer que essas espécies sejam reprodutivamente isoladas umas das outras. De fato, definimos uma espécie como uma população de organismos que trocam genes entre si e que são reprodutivamente isolados de outras populações. Os mecanismos de isolamento reprodutivo podem ser pré- ou pós-zigóticos. Os mecanismos de isolamento pré-zigóticos são os que impedem a união de gametas de duas espécies. Esses mecanismos podem ser de incompatibilidade comportamental dos machos e fêmeas de espécies diferentes, diferenças na época ou local de sua atividade sexual, diferenças anatômicas que tornam impossível mecanicamente a reprodução, ou incompatibilidade fisiológica dos próprios gametas. Mecanismos de isolamento pós-zigótico incluem a incapacidade de os embriões híbridos se desenvolverem até a vida adulta, a esterilidade dos adultos híbridos e a degradação das gerações posteriores de genótipos recombinantes. Na maior parte, as diferenças genéticas responsáveis pelo isolamento entre espécies proximamente relacionadas espalham-se por todos os cromossomos, embora em espécies com determinação cromossômica sexual possa haver uma concentração de genes de incompatibilidade no cromossomo sexual.

No geral, a evolução genética é um processo histórico que está sujeito a contingências históricas e ao acaso, mas é restrita pela necessidade dos organismos de sobreviver e se reproduzir em um mundo em constante mudança.

► Termos-chave

característica canalizada	isolamento pré-zigótico	seleção direcional
diversificação	mapa de sintenia	seleção natural
efeito do fundador	melanina	seleção purificadora
epistasia de sinal	melanismo	substituição não-sinônima
especiação alopátrica	panorama adaptativo	substituição sinônima
espécie	populações alopátricas	superfície adaptativa
evolução filética	relógio molecular	
isolamento pós-zigótico	seleção balanceadora	

► Problemas Resolvidos

Problema resolvido 1. Um entomologista que estuda insetos que se alimentam de vegetação apodrecida descobriu um caso interessante de diversificação de borrachudos de fungo em várias ilhas em um arquipélago. Cada ilha tem uma população de borrachudos de morfologia extremamente similar, embora eles não sejam idênticos aos das outras ilhas, mas cada um vive em um tipo diferente de vegetação apodrecida que não está presente nas outras ilhas. O entomologista postula que essas populações são espécies proximamente relacionadas que divergiram adaptando-se a se alimentar em condições de apodrecimento levemente diferentes.

Para apoiar sua hipótese, ele efetua um estudo eletroforético da enzima álcool desidrogenase em populações diferentes. Ele descobre que cada população é caracterizada por uma forma eletroforeticamente diferente de álcool desidrogenase, e então raciocina que cada uma dessas formas de álcool desidrogenase é especificamente adaptada a álcoois particulares que são produzidos na fermentação da vegetação característica de determinada ilha. Além disso, existe algum polimorfismo de álcool desidrogenase dentro de cada ilha, mas a frequência de alelos variantes é baixa em cada ilha e pode ser facilmente explicada como resultado de uma mutação ocasional ou raras migrações de outras ilhas. Este borrachudo de fungo tornou-se então um exemplo de livros de como a diversidade de espécies pode surgir por seleção natural adaptando cada espécie recém-formada a um ambiente diferente.

Uma cética geneticista de populações leu sobre o caso em um livro e imediatamente teve algumas dúvidas. Pareceu a ela que, em vista da evidência, uma explicação igualmente plausível é que essas populações de borrachudos não são espécies, mas sim raças geográficas locais que se tornaram levemente diferenciadas morfologicamente por deriva genética aleatória. Além disso, as formas eletroforeticamente diferentes da proteína álcool desidrogenase podem ser variantes fisiologicamente equivalentes de um gene sofrendo evolução molecular neutra em populações isoladas.

Imagine um programa de pesquisa que possa distinguir essas explicações alternativas. Como você testaria se as populações diferentes são de fato espécies diferentes?

Como você poderia testar a hipótese de que formas diferentes de álcool desidrogenase divergiram seletivamente?

SOLUÇÃO

Para testar a distinção das espécies dos borrachudos diferentes, é necessário ser capaz de manipulá-los e cultivá-los em cativeiro. Se eles não podem ser cultivados no laboratório ou em estufas, a distinção das espécies não pode ser estabelecida. A compatibilidade do comportamento reprodutivo das formas diferentes pode ser testada colocando-se uma mistura de machos de duas populações diferentes com fêmeas de uma das formas para ver se existe alguma preferência reprodutiva das fêmeas. O mesmo experimento pode então ser repetido com fêmeas misturadas e machos de uma forma, e com misturas de machos e fêmeas de ambas as formas. A partir de tais experimentos, podem ser observados padrões de preferência reprodutiva. Mesmo que exista alguma pequena quantidade de formas reprodutivas diferentes, elas só podem ocorrer devido a condições não-naturais nas quais está sendo feito o teste. Por outro lado, não deverá ocorrer nenhum tipo de reprodução, mesmo entre as mesmas formas, porque faltam os aspectos necessários para a reprodução, em cujo caso nada pode ser concluído.

Se ocorrerem reproduções entre formas diferentes, a sobrevivência dos híbridos interpopulacionais pode ser comparada com as reproduções intrapopulacionais. Se os híbridos sobreviverem, sua fertilidade pode ser testada tentando retrocruzá-los com duas linhagens parentais diferentes. Como nos testes de reprodução, sob as condições não-naturais do laboratório ou da estufa, é possível alguma sobrevivência ou fertilidade dos híbridos das espécies, muito embora o isolamento na natureza seja completo. Qualquer redução clara na sobrevivência ou fertilidade observada dos híbridos é uma forte suposição de evidência de que pertençam a espécies diferentes.

Para testar se as diferentes seqüências de aminoácidos subjacentes às diferenças de mobilidade eletroforética são o resultado de divergência seletiva, é necessário um programa de seqüenciamento de DNA do locus de álcool desidrogenase. As amostras replicadas de seqüências de *Adh* de cada uma das populações de ilhas devem ser obtidas. O número necessário de tais seqüências de cada população depende do grau de polimorfismo de nucleotídeos que está presente nas

populações, mas os resultados de muitos loci em muitas espécies sugere que, como regra geral, pelo menos sejam obtidas 10 seqüências de cada população. Os sítios polimórficos dentro das populações são classificados em sítios não-sinônimos (*a*) e sinônimos (*b*). As diferenças fixas de nucleotídeos entre populações são classificadas em diferenças não-sinônimas (*c*) e sinônimas (*d*). Se a divergência entre as populações é puramente o resultado de deriva genética aleatória, então esperamos que a/b seja igual a c/d . Se, por outro lado, houve divergência seletiva, deve haver um excesso de diferenças fixadas não-sinônimas, e, logo, a/b deve ser menor que c/d . A igualdade dessas proporções pode ser testada por um teste de χ^2 de contingência 2×2 da forma

	Polimorfismos	
	Não-sinônimos	Sinônimos
Diferenças	<i>a</i>	<i>b</i>
populacionais	<i>c</i>	<i>d</i>

$$\chi^2 = \frac{(a + b + c + d) (ad - bc)^2}{(a + c) (b + d) (a + b) (c + d)}$$

Problema resolvido 2. Duas espécies proximamente relacionadas são encontradas como sendo fixadas por dois alelos diferentes detectados eletroforéticamente em um locus codificando uma enzima. Como você pode demonstrar que essa diferença é um resultado de seleção natural e não de evolução neutra?

SOLUÇÃO

- Obtenha seqüências de DNA do gene de um número de indivíduos separados ou linhagens de cada uma das duas espécies. Dez ou mais seqüências de cada espécie seriam desejáveis.
- Tabule as diferenças de nucleotídeos entre indivíduos dentro de cada espécie (polimorfismos) e classifique essas diferenças como as que resultam em mudanças de aminoácidos (polimorfismos de substituição) ou as que não mudam o aminoácido (polimorfismos sinônimos).

c. Faça a mesma tabulação de mudanças de substituição e sinônimas para as diferenças entre as espécies, contando apenas as diferenças que diferenciam totalmente as espécies. Isto é, não conte um polimorfismo em uma espécie que inclua uma variante que é vista na outra espécie.

d. Se a proporção de diferenças de substituição entre as espécies por diferenças sinônimas entre as espécies for maior do que a proporção de polimorfismos de substituição por polimorfismos sinônimos, então selecione a mudança de aminoácidos.

e. Teste o significado estatístico da maior proporção observada por um teste de χ^2 2×2 do seguinte quadro:

		Polimorfismos	
		Substituição	Sinônimos
Diferenças de	Substituição	<i>a</i>	<i>b</i>
espécies	Sinônimos	<i>c</i>	<i>d</i>

$$\chi^2 = \frac{(a + b + c + d) (ad - bc)^2}{(a + c) (b + d) (a + b) (c + d)}$$

Problema resolvido 3. Como pode a evolução molecular de um conjunto de proteínas diferentes ser usado para dar evidências da importância relativa da seqüência exata de aminoácidos para a função de cada proteína?

SOLUÇÃO

Obtenha seqüências de DNA dos genes de cada proteína de uma grande variedade de espécies divergentes cujo tempo aproximado a um ancestral comum é conhecido pelo registro fóssil. Traduza as seqüências de DNA em seqüências de aminoácidos. Para cada proteína, plote a diferença observada de aminoácidos para cada par de espécies contra o tempo estimado de divergência dessas espécies. A linha para cada proteína terá uma inclinação que é proporcional à intensidade de restrição funcional na substituição de aminoácidos nessa proteína. Proteínas altamente restritas terão taxas muito baixas de substituição, enquanto as proteínas mais tolerantes terão inclinações maiores.

► Problemas

PROBLEMAS BÁSICOS

- Qual a diferença entre uma transformação e um esquema variacional de evolução? Cite um exemplo de cada (não incluindo a teoria de Darwin da evolução orgânica).
- Quais são os três princípios da teoria de Darwin da evolução variacional?
- Por que a explicação mendeliana da herança é essencial para o mecanismo variacional de Darwin para a evolução? Quais seriam as conseqüências para a evolução se a herança fosse por mistura de sangue? Qual seria a conseqüência para a evolução se os heterozigotos não tivessem uma segregação de exatamente 50% de cada um dos dois alelos em um locus, mas fossem consistentemente desviados para um ou outro alelo?

- O que é uma raça geográfica? Qual a diferença entre raça geográfica e uma espécie separada? Sob que condições as raças geográficas de uma espécie tornam-se novas espécies?

PROBLEMAS DESAFIADORES

- Se a taxa de mutação para um novo alelo é de 10^{-5} , que tamanho as populações isoladas devem ter para evitar a diferenciação casual entre elas na freqüência desse alelo?
- Suponha que um número de populações locais de uma espécie tenha cada uma cerca de 10.000 indivíduos e que não ocorra migração entre elas. Suponha ainda que elas foram originalmente estabelecidas de uma grande população com a freqüência de um alelo *A* em algum locus

igual a 0,4. Mostre por desenhos aproximados qual seria a distribuição de frequências de alelos entre as populações locais após 100, 1.000, 5.000, 10.000 e 100.000 gerações de isolamento.

7. Mostre os resultados para as populações descritas no Problema 6 se houver uma troca de migrantes entre as populações na taxa de (a) um indivíduo migrante por população a cada 10 gerações; (b) um indivíduo migrante por população a cada geração.
8. Suponha que uma população esteja segregando dois alelos em cada um de dois loci e que as probabilidades relativas de sobrevivência até a maturidade sexual dos zigos de nove genótipos sejam as seguintes:

	A/A	A/a	a/a
B/B	0,95	0,90	0,80
B/b	0,90	0,85	0,70
b/b	0,90	0,80	0,65

Calcule a adaptabilidade média, \bar{W} , da população se as frequências de alelos são $p(A) = 0,8$ e $p(B) = 0,9$. Que direção de mudança você espera em frequências alélicas na próxima geração? Faça o mesmo cálculo e previsão para as frequências de alelos $p(A) = 0,2$ e $p(B) = 0,2$. Pela inspeção das adaptabilidades genotípicas, quantos picos adaptativos existem? Quais são as frequências alélicas nos picos?

9. Suponha que as adaptabilidades genotípicas no Problema 8 fossem

	A/A	A/a	a/a
B/B	0,9	0,8	0,9
B/b	0,7	0,9	0,7
b/b	0,9	0,8	0,9

Calcule a adaptabilidade média, \bar{W} , para as frequências alélicas $p(A) = 0,5$ e $p(B) = 0,5$. Que direção de mudança você espera para as frequências alélicas na geração seguinte? Repita o cálculo e a previsão para $p(A) = 0,1$ e $p(B) = 0,1$. Pela inspeção das adaptabilidades dos genótipos, quantos picos adaptativos existem e onde eles estão situados?

10. O gene *MC1R* afeta a cor da pele e dos cabelos em humanos. Existem pelo menos 13 polimorfismos do gene em populações europeias e asiáticas, 10 dos quais não são sinônimos. Em africanos, existem pelo menos 5 polimorfismos do gene, nenhum dos quais são sinônimos. Qual pode ser uma explicação para as diferenças na variação de *MC1R* entre africanos e não-africanos?
11. As proteínas opsina detectam a luz nas células fotorreceptoras do olho e são necessárias para a visão de cores. O macaco-coruja, o *bush baby* e o rato cego subterrâneo

têm mutações diferentes em um gene de opsina que o tornam não-funcional. Explique o porquê disso.

12. A falta total ou parcial de membros evoluiu muitas vezes nos vertebrados (cobras, lagartos, manatim, baleias). Você espera que as mutações na evolução da falta de membros sejam nas seqüências codificantes ou não-codificantes dos genes ferramentais (*toolkit*)? Por quê?
13. Várias espécies de *Drosophila* com asas sem manchas descendem de um ancestral com manchas. Você preveria que a perda da formação de manchas codifica ou não codifica mudanças nos genes de pigmentação? Como você testaria qual é o caso?
14. Qual a evidência molecular de que a seleção natural inclui a “rejeição de mudança prejudicial”?
15. Qual a evidência de que a formação de poliplóides foi importante na evolução de plantas?
16. Qual a evidência de que a duplicação gênica foi a fonte das famílias de genes α e β na hemoglobina humana?
17. O alelo de grupo sanguíneo humano I^B tem uma frequência de cerca de 0,10 em populações europeias e asiáticas, mas é quase totalmente ausente em populações de nativos americanos. Que explicações justificam essa diferença?
18. *Drosophila pseudoobscura* e *D. persimilis* hoje são consideradas espécies separadas, mas originalmente eram classificadas como raça A e raça B de uma única espécie. Elas são morfologicamente indistinguíveis uma da outra, exceto por uma pequena diferença na genitália dos machos. Quando cruzadas no laboratório, é produzida uma prole F_1 abundante de ambos os sexos. Crie um programa de observações e experimentos para testar a afirmativa de que as duas formas são espécies diferentes.
19. Usando os dados de similaridade de aminoácidos das cadeias de globina α , β , γ , ζ e ϵ dados no Quadro 19.4, desenhe uma árvore ramificada da evolução dessas cadeias a partir de uma seqüência ancestral original na qual a ordem de ramificação no tempo seja o mais consistente possível com a similaridade observada de aminoácidos na suposição de um relógio molecular.
20. Os estudos de seqüenciamento de DNA para um gene em duas espécies proximalmente relacionadas produziram os seguintes números de sítios que variam:

Polimorfismos sinônimos	50
Espécies diferentes não-sinônimas	2
Espécies diferentes sinônimas	18
Polimorfismos não-sinônimos	20

Esse resultado apóia a evolução neutra do gene? Ela apóia uma substituição adaptativa de aminoácidos? Que explicação você ofereceria para essas observações?

EXPLORANDO GENOMAS Um Tutorial de Bioinformática Baseado na Web

Medindo a Distância Filogenética

As seqüências de dados nos permitem estimar as distâncias evolutivas entre organismos com base na extensão da divergência

de seqüência. No tutorial Genômica em www.whfreeman.com/iga9e, são usadas comparações de seqüências para gerar ou apoiar nossas conclusões de acordo com a estrutura da árvore evolutiva.

Isolamento e Manipulação de Gene



Injeção de DNA exógeno em uma célula animal. A microagulha usada para a injeção é mostrada à direita, e a pipeta que segura uma célula é mostrada à esquerda. [Copyright M. Baret/Rapho/Photo Researchers.]

As plantas e animais domesticados são os produtos de seleção artificial pelos humanos para características desejáveis encontradas em populações selvagens. A revolução agrícola, que ocorre há mais de 10.000 anos, foi desencadeada pelos primeiros fazendeiros que escolheram colher sementes de plantas capazes de expressar certas características que acentuariam seu valor como fonte alimentar. Um exemplo de tal característica na domesticação de várias espécies de plantas é a retenção de sementes. As sementes de plantas selvagens geralmente caem das plantas (dispersam), espalhando assim a prole a distância do genitor. Ao contrário, as sementes da maioria dos cultivos não se dispersam e são mais facilmente colhidas. Não é difícil imaginar um cenário no qual uma tribo antiga pôs-se a colher sementes para alimento e chegou até um indivíduo ou população que tinha suas sementes. Não só essas sementes seriam comidas, mas também seriam plantadas.

A dispersão é uma das características que transformam uma espécie selvagem em cultivos domesticados. Às vezes, essa transformação foi tão intensa

20

Perguntas Fundamentais

- Como um gene é isolado e amplificado por clonagem?
- Como DNA ou RNA específicos são identificados em misturas?
- Como o DNA é amplificado sem clonagem?
- Como o DNA amplificado é usado em genética?
- Como as tecnologias de DNA são aplicadas à medicina?

Tópicos

- 20.1 Gerando moléculas de DNA recombinante
- 20.2 Amplificação de DNA *in vitro*: a reação em cadeia da polimerase
- 20.3 Determinando a sequência de bases de um segmento de DNA
- 20.4 Análise genética direta usando clonagem posicional
- 20.5 Detecção de alelos de doenças humanas: diagnóstico genético molecular
- 20.6 Engenharia genética

A genitora do milho domesticado



FIG. 20.1 Uma subespécie de teosinto, *Zea mays* ssp. *parviglumis*, cresce em uma ravina perto de Teloloapan na drenagem do rio Balsas, Guerrero, México. [Cortesia de Hugh Iltis.]

que é difícil identificar o genitor de determinado cultivo. Um exemplo é o milho, que é uma gramínea mas não se assemelha a nenhuma espécie de gramínea. Em 1939, George Beadle propôs que, a despeito de grandes diferenças morfológicas, o genitor mais provável de milho é a gramínea teosinto, que ainda cresce em estado selvagem no México (Figura 20.1). Como os antigos fazendeiros foram capazes de fazer a transformação radical de teosinto em milho? Como veremos, usando técnicas descritas neste capítulo, os geneticistas isolaram um dos genes mais importantes nessa transformação.

Genes, tais como os selecionados por humanos na domesticação de plantas e animais, são o foco central da Genética, e, assim, claramente é desejável ser capaz de isolar um gene de interesse (ou qualquer região do DNA) do genoma e amplificá-lo para obter uma quantidade suficiente para estudo. **Tecnologia do DNA** é um termo que descreve as técnicas coletivas para obter, amplificar e manipular fragmentos específicos de DNA. Desde a metade dos anos de 1970, o desenvolvimento da tecnologia do DNA revolucionou o campo da Biologia, abrindo muitas áreas de pesquisa para a investigação molecular. A **Engenharia Genética**, a aplicação da tecnologia do DNA a problemas específicos biológicos, médicos ou agrícolas, hoje é um ramo bem estabelecido da tecnologia. A **Genômica** é a extensão da tecnologia à análise global dos ácidos nucleicos presentes em um núcleo, uma célula, um organismo ou um grupo de espécies correlatas (veja Capítulo 13). Mais adiante neste capítulo, veremos como as técnicas da tecnologia do DNA e genômica, junta-

mente com os métodos apresentados nos Capítulos 2 e 4, podem ser usados para isolar e identificar um gene.

► 20.1 Gerando Moléculas de DNA Recombinante

Como as amostras de segmentos individuais de DNA podem ser isoladas? Essa tarefa pode inicialmente assemelhar-se a procurar uma agulha no palheiro. Um *insight* crucial foi que os pesquisadores puderam criar grandes amostras de DNA que eles precisavam enganando a maquinaria de replicação do DNA para replicar o segmento de DNA em questão. Tal replicação pode ser feita ou dentro de células bacterianas vivas (*in vivo*), ou em um tubo de ensaio (*in vitro*). Exploraremos o enfoque *in vitro* na Seção 20.2.

No enfoque *in vivo*, um pesquisador começa com uma amostra de moléculas de DNA contendo o gene de interesse. Essa amostra é chamada de **DNA doador** e, mais frequentemente, é um genoma inteiro. Os fragmentos de DNA doador são inseridos nos cromossomos “acessórios” não-essenciais (tais como plasmídeos ou vírus bacterianos modificados). Esses cromossomos acessórios “levarão” e amplificarão o gene de interesse, sendo assim chamados de **vetores**. Primeiro, as longas moléculas doadoras de DNA são cortadas em centenas ou milhares de fragmentos de tamanho mais manejável. Em seguida, cada fragmento é fundido com um cromossomo vetor para formar moléculas de **DNA recombinante**. As moléculas de DNA recombinante são inseridas em células bacterianas, e, geralmente, apenas uma molécula recombinante é captada por cada célula. Como o cromossomo acessório normalmente é amplificado por replicação, a molécula recombinante é similarmente amplificada durante o crescimento e divisão da célula bacteriana na qual o cromossomo reside. Esse processo resulta em um *clone* de células idênticas, cada uma contendo a molécula de DNA recombinante, e, assim, essa técnica de amplificação é chamada de **clonagem de DNA**. O estágio seguinte é encontrar o clone raro contendo o DNA de interesse.

Para ilustrar como é feito o DNA recombinante, consideremos a clonagem do gene humano de insulina, um hormônio proteico usado no tratamento do diabetes. O diabetes é uma doença na qual os níveis de açúcar no sangue são anormalmente altos seja porque o corpo não produz insulina suficiente (diabetes tipo I), seja porque as células são incapazes de responder à insulina (diabetes tipo II). Nas formas brandas do tipo I, o diabetes pode ser tratado com restrições dietéticas, mas, para muitos pacientes, os tratamentos diários com insulina são necessários. Até cerca de 20 anos atrás, as vacas eram a principal fonte da proteína insulina. A proteína era colhida do pâncreas dos animais sacrificados em matadouros e purificada em grande escala para eliminar a maioria das proteínas e outros contaminantes dos extratos de pâncreas. Então, em 1982, a primeira insulina humana recombinante surgiu no mercado. A insulina humana podia ser feita em forma pura, a um custo menor, e em escala industrial, pois foi produzida em bactérias por técnicas de DNA recombinante. A insulina recombinante é uma proporção mais alta

das proteínas na bactéria do que a insulina bovina; assim, a purificação da proteína é mais fácil. Seguiremos as etapas gerais necessárias para fazer qualquer DNA recombinante e aplicá-las à insulina.

Tipo de DNA doador

A escolha do DNA a ser usado como doador pode parecer óbvia, mas na verdade existem três possibilidades.

- **DNA genômico.** Esse DNA é obtido diretamente dos cromossomos do organismo em estudo. É a fonte mais direta de DNA. O DNA cromossômico precisa ser cortado antes que a clonagem seja possível.
- **cDNA. DNA complementar (cDNA)** é uma versão bifilamentar do DNA de uma molécula de mRNA. Nos eucariotes superiores, um mRNA é um previsor mais útil de uma sequência polipeptídica do que é a sequência genômica, pois os íntrons foram removidos. Os pesquisadores preferem usar um cDNA do que o próprio mRNA, pois os RNA são inerentemente menos estáveis do que o DNA e não existem técnicas para, rotineiramente, amplificar e purificar moléculas individuais de RNA. O cDNA é feito de mRNA com o uso de uma enzima especial chamada *transcriptase reversa*, originalmente isolada de retrovírus. Usando uma molécula de mRNA como molde, a transcriptase reversa sintetiza uma molécula de DNA unifilamentar que pode ser usada como molde para a síntese de DNA bifilamentar. O cDNA não precisa ser cortado para ser clonado.
- **DNA sintetizado quimicamente.** Às vezes, um pesquisador precisa incluir em uma molécula de DNA recombinante uma sequência específica que, por algum motivo, não pode ser isolada do DNA ou cDNA genômico natural disponível. Se a sequência de DNA é conhecida (em geral, de uma sequência genômica completa), então o gene pode ser sintetizado quimicamente usando técnicas automatizadas.

Para criar bactérias que expressam insulina humana, o cDNA foi a escolha inicial porque as bactérias não têm a habilidade de remover os íntrons presentes no DNA genômico natural. Além disso, devido a um fenômeno chamado de **tendenciosidade de códon**, o cDNA humano tinha que ser modificado quimicamente, de modo a poder ser otimamente expresso em bactérias. A tendenciosidade de códon vem da redundância do código genético, o fato de que alguns aminoácidos são especificados por múltiplos códons (veja Figura 9.6). Por exemplo, o código genético inclui seis códons de leucina e quatro códons de serina. A tendenciosidade refere-se a uma preferência do organismo por um ou mais códons para um aminoácido particular. Para fazer grandes quantidades de insulina humana, o cDNA humano tinha que ser modificado quimicamente para corresponder à tendenciosidade de códon de *E. coli*.

Cortando o DNA genômico

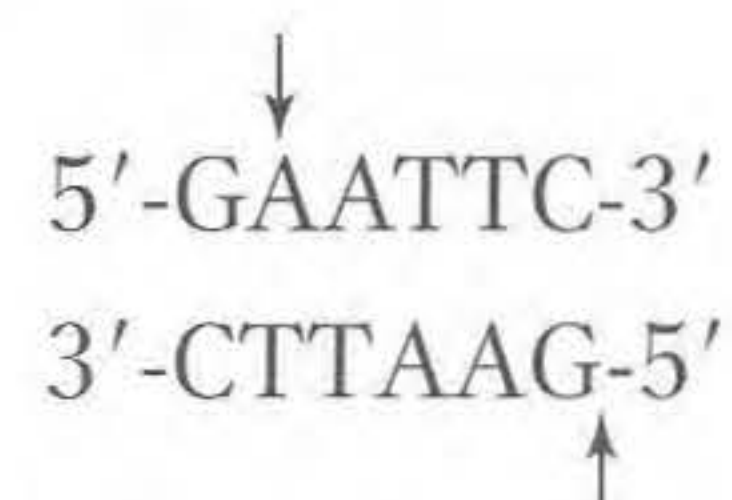
As grandes moléculas de DNA de tamanho cromossômico devem ser cortadas em fragmentos de tamanho muito menor.

A maioria dos cortes são feitos com o uso de **enzimas de restrição**. Essas enzimas cortam sequências-alvo específicas de DNA, chamadas de *sítios de restrição*, e essa propriedade é uma das características principais que tornam as enzimas de restrição adequadas para a manipulação do DNA. Puramente por acaso, qualquer molécula de DNA, seja ela derivada de vírus, mosca ou humanos, contém sítios-alvo de enzimas de restrição. Assim, uma enzima de restrição cortará o DNA em um conjunto de **fragmentos de restrição** determinados pelos locais dos sítios de restrição.

Outra propriedade-chave de algumas enzimas de restrição é que elas fazem “pontas adesivas”. Vejamos um exemplo. A enzima de restrição *EcoRI* (de *E. coli*) reconhece a seguinte sequência de seis pares de nucleotídeos no DNA de qualquer organismo:



Esse tipo de segmento é chamado de um **palíndromo de DNA**, o que significa que ambos os filamentos têm a mesma sequência de nucleotídeos mas com orientação em polaridade inversa. Diferentes enzimas de restrição cortam em sequências palindrômicas diferentes. Às vezes, os cortes são na mesma posição em cada um dos filamentos de polaridade inversa. Entretanto, as enzimas de restrição mais úteis fazem cortes que são desencontrados. Por exemplo, a enzima *EcoRI* faz cortes apenas entre os nucleotídeos G e A em cada filamento do palíndromo.



Esses cortes desencontrados deixam um par de pontas adesivas idênticas, cada uma com um único filamento de cinco bases de tamanho. As pontas são chamadas de *adesivas* porque, sendo unifilamentares, elas podem fazer par de bases (*i.e.*, grudar) com uma sequência complementar. O pareamento unifilamentar desse tipo às vezes é chamado de **hibridização**. A Figura 20.2 (em cima, à esquerda) ilustra a enzima de restrição *EcoRI* fazendo um único corte em uma molécula circular de DNA tal como um plasmídeo; o corte abre o círculo, e a molécula linear resultante tem duas pontas adesivas. Elas podem então se hibridizar com um fragmento de uma molécula diferente de DNA que tenha as mesmas pontas adesivas complementares.

Enzimas de restrição com sequências de especificidades diferentes hoje são conhecidas. Algumas enzimas, tais como *EcoRI* ou *PstI*, fazem cortes desencontrados, enquanto outras, tais como *SmaI*, fazem cortes que deixam pontas rombudas. Mesmo os cortes rombudos, que não têm pontas adesivas, podem ser usados para fazer DNA recombinante.

As enzimas de restrição cortam o DNA em fragmentos de tamanho manejável, e muitas delas geram pontas adesivas unifilamentares adequadas para fazer DNA recombinante.

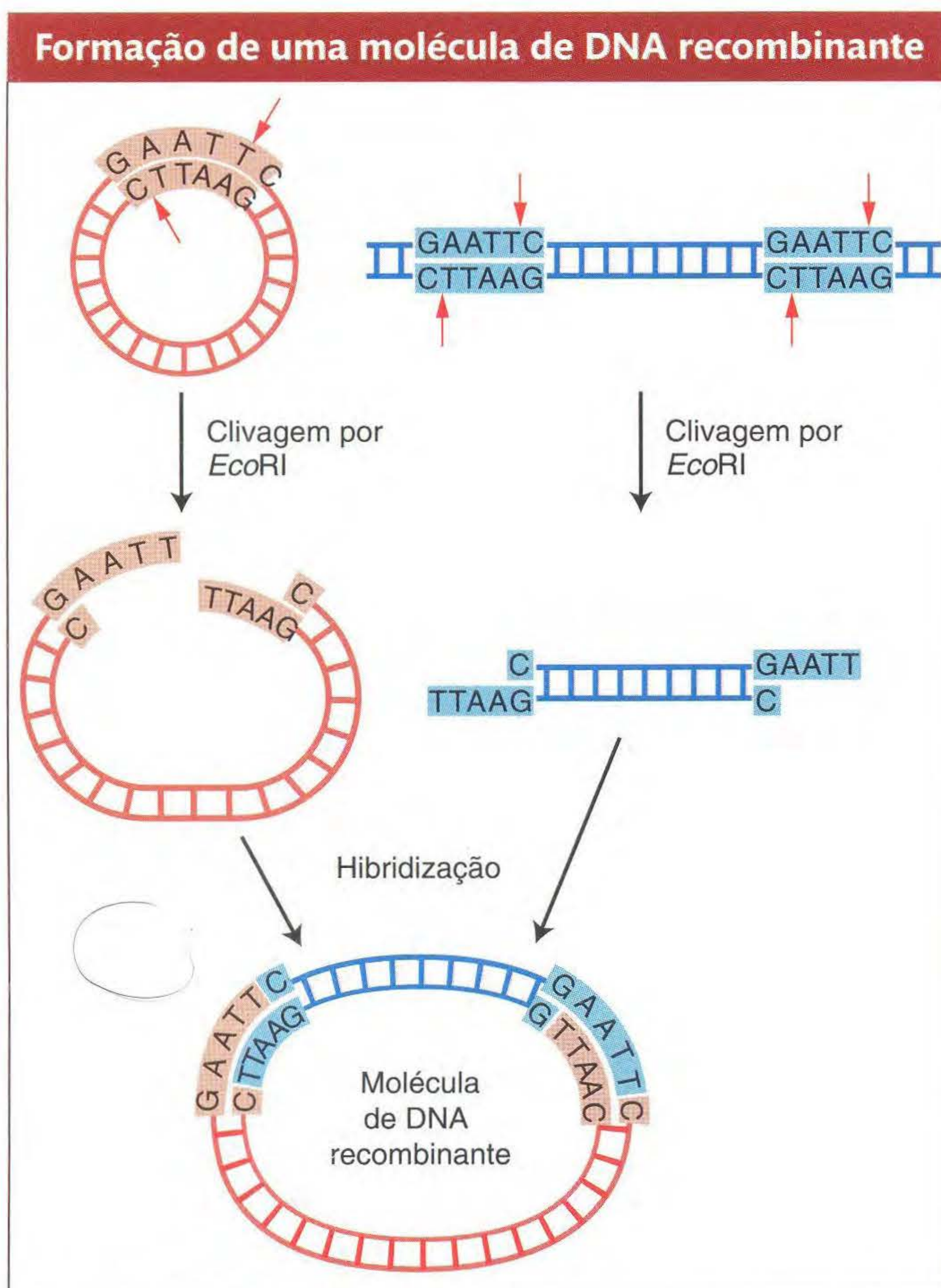


FIG. 20.2 Enzimas de restrição são usadas para fazer DNA recombinante. A enzima de restrição *EcoRI* corta uma molécula circular de DNA que tem uma sequência-alvo, resultando em uma molécula linear com pontas adesivas unilaterais. Devido à complementariedade, outras moléculas lineares com pontas adesivas cortadas com *EcoRI* podem se hibridizar ao DNA circular linearizado, formando uma molécula de DNA recombinante.

Ligação do DNA doador e vetor

Mais comumente, tanto o DNA doador quanto o vetor são digeridos por uma enzima de restrição que produz pontas adesivas complementares e são misturados em um tubo de ensaio para permitir que as pontas adesivas do DNA doador e vetor se liguem para formar moléculas recombinantes. A Figura 20.3 mostra um DNA plasmidial bacteriano que leva um único sítio de restrição *EcoRI*; assim, a digestão com a enzima de restrição *EcoRI* converte o DNA circular em uma única molécula linear com pontas adesivas. O DNA doador de qualquer outra fonte, tal como o DNA humano, também é tratado com a enzima *EcoRI* para produzir uma população de fragmentos levando as mesmas pontas adesivas. Quando duas populações são misturadas sob condições fisiológicas apropriadas, os fragmentos de DNA de duas fontes podem se hibridizar, pois as duplas hélices formam-se entre suas pontas adesivas (Figura 20.3b). Existem muitas moléculas de plasmídeo abertas na solução, bem como muitos fragmentos diferentes de *EcoRI* do DNA doador. Portanto, um arranjo

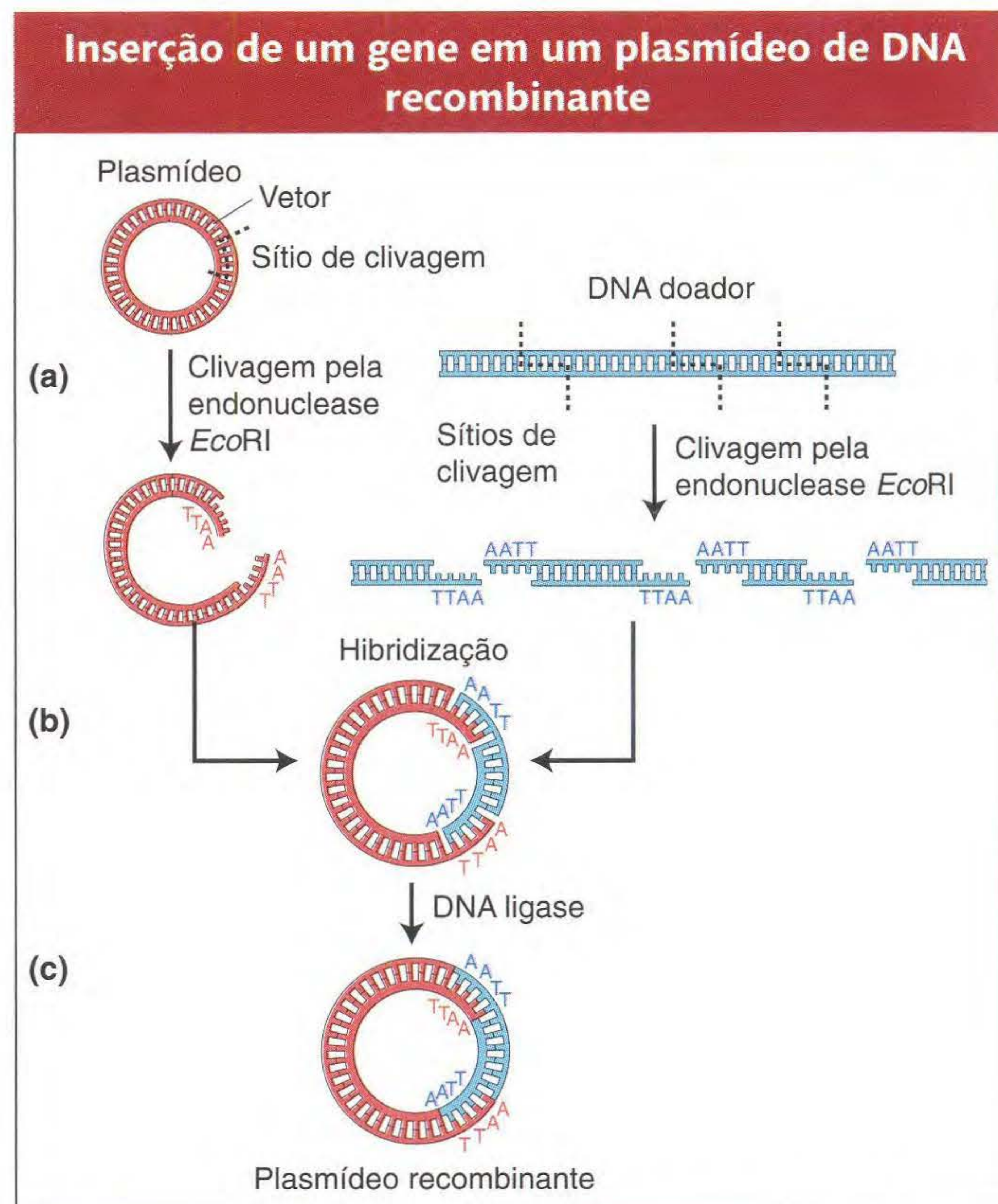


FIG. 20.3 Método para gerar um DNA plasmidial recombinante contendo genes derivados de DNA doador. [De S.N. Cohen, "The Manipulation of Genes". Copyright 1975 by Scientific American, Inc. Todos os direitos reservados.]

diferente de plasmídeos recombinados com fragmentos doadores diferentes será produzido. Nesse estágio, as moléculas hibridizadas não uniram covalentemente os arcabouços açúcar-fosfato. Entretanto, os arcabouços podem ser unidos pela adição da enzima **DNA ligase**, que cria ligações fosfodiéster nas junções (Figura 20.3c).

O DNA complementar pode ser unido ao vetor com o uso apenas da ligase. Alternativamente, podem ser adicionadas pontas adesivas curtas a cada ponta de um plasmídeo e vetor.

Outra consideração nesse estágio é que, se o gene clonado vai ser transcrito e traduzido no hospedeiro bacteriano, ele deve ser inserido perto de sequências reguladoras bacterianas. Assim, para ser capaz de produzir insulina humana em bactérias, o gene deve estar adjacente a sequências reguladoras bacterianas corretas.

Amplificação dentro de uma bactéria

A amplificação aproveita-se de processos genéticos procarióticos, incluindo os de transformação bacteriana, replicação de plasmídeos e crescimento de bacteriófagos, todos discutidos no Capítulo 5. A Figura 20.4 ilustra a clonagem de um segmento de DNA doador. Um único vetor recombinante entra em uma célula bacteriana e é amplificado pela replicação que ocorre na divisão celular. Existem geralmente muitas cópias de cada vetor em cada bactéria. Assim, após a amplifi-

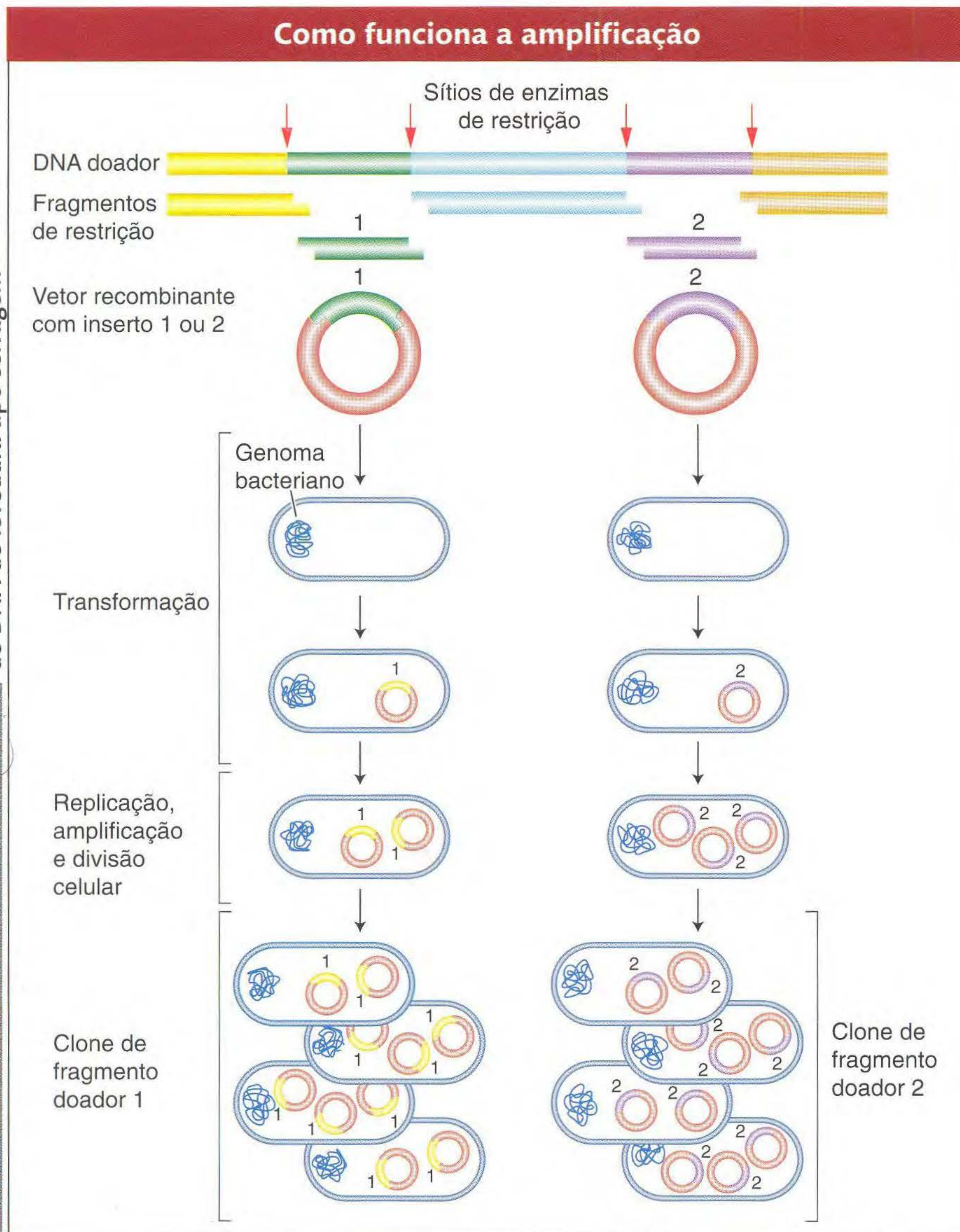


FIG. 20.4 A estratégia geral usada para clonar um gene. O tratamento com enzima de restrição do DNA doador e vetor permite a inserção de fragmentos únicos em vetores. Um único vetor entra na bactéria hospedeira, onde a replicação e a divisão celular resultam em um grande número de cópias do fragmento doador.

cação, uma colônia de bactérias tipicamente conterá bilhões de cópias do único inserto de DNA doador fundido a seu cromossomo acessório. Esse conjunto de cópias amplificadas de um único fragmento de DNA doador dentro do vetor de clonagem é o *clone* de DNA recombinante. A replicação de moléculas recombinantes explora os mecanismos normais que a bactéria usa para replicar o DNA cromossômico. Um requisito básico é a presença de uma origem de replicação de DNA (como descrito no Capítulo 7).

Escolha dos vetores de clonagem Os vetores devem ser moléculas pequenas para a conveniente manipulação. Eles devem ser capazes de replicação prolífica em uma célula viva de modo a amplificar o fragmento doador inserido. Ela também deve ter convenientes sítios de restrição nos quais o DNA a ser clonado pode ser inserido. Idealmente, o sítio de restrição deve estar presente apenas uma vez no vetor porque então os fragmentos de restrição do DNA doador irão inserir-se apenas nesse único local no vetor. Também

é importante ter um modo para identificar e recuperar rapidamente a molécula recombinante. Vários vetores de clonagem estão em uso hoje em dia, adequados aos diferentes tamanhos de inserto de DNA ou usos diferentes do clone. Algumas classes gerais de vetores de clonagem são citadas a seguir.

Vetores plasmidiais Como já descrito, os plasmídeos bacterianos são pequenas moléculas circulares de DNA que replicam o seu DNA independentemente do cromossomo bacteriano. Os plasmídeos rotineiramente usados como vetores são os que levam genes para resistência a drogas. Esses genes de resistência a drogas fornecem um modo conveniente para selecionar células transformadas por plasmídeos: as células ainda vivas após a exposição a uma droga devem levar os vetores plasmidiais contendo o inserto de DNA. A Figura 20.5 ilustra como as características de um vetor, chamado de pUC18, permitem a detecção de plasmídeos que contêm DNA inserido.

Um plasmídeo vetor, pUC18

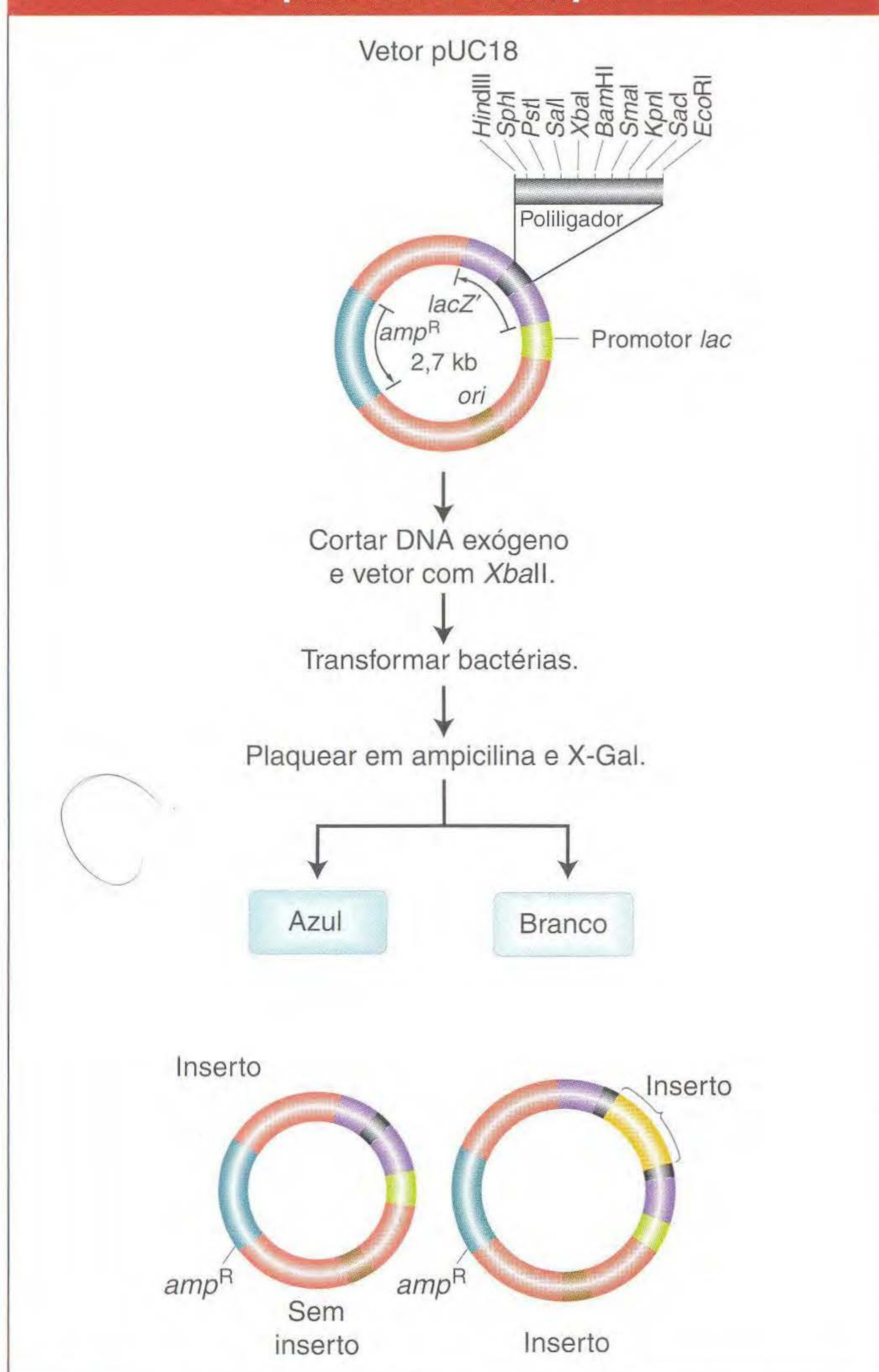


FIG. 20.5 Um plasmídeo vetor, pUC18, foi criado para uso como um vetor para clonagem de DNA. A inserção em pUC18 é detectada pela inativação da função de β -galactosidase de *lacZ'*, resultando na incapacidade em converter o substrato artificial X-Gal em um corante azul. O poliligador tem vários sítios de restrição alternativos nos quais o DNA doador pode ser inserido.

Vetores bacteriófagos Diferentes classes de vetores bacteriófagos podem levar tamanhos diferentes de inserto de DNA doador. Determinado bacteriófago pode abrigar uma quantidade padrão de DNA como um inserto “embalado” dentro de uma partícula de fago. O bacteriófago λ (lambda) é um efetivo vetor de clonagem para insertos de DNA bifilamentar de até cerca de 15 kb. As cabeças do fago lambda podem acomodar moléculas de DNA não maiores do que cerca de 50 kb de tamanho (o tamanho de um cromossomo λ normal). A parte central do genoma do fago não é necessária para a replicação ou embalagem de moléculas de DNA λ em *E. coli* e, assim, pode ser removida usando enzimas de restrição e sendo descartada. A parte central deletada é então substituída por insertos de DNA doador. Um inserto terá de 10 a 15 kb de tamanho, porque um inserto dessa

dimensão leva o tamanho cromossômico total a ter de novo seus 50 kb normais.

Como a Figura 20.6 mostra, as moléculas recombinantes podem ser diretamente embaladas nas cabeças de fago *in vitro* e, então, introduzidas na bactéria. Alternativamente, as moléculas recombinantes podem ser diretamente transformadas em *E. coli*. Em ambos os casos, a presença de uma placa de fago na camada de bactérias automaticamente indica a presença de fagos recombinantes portando um inserto.

Vetores para insertos maiores de DNA O plasmídeo padrão e os vetores fago λ descritos podem aceitar DNA doador de tamanhos tão grandes quanto 25 a 30 kb. Entretanto, muitos experimentos requerem insertos além desse limite superior. Para atender a essas necessidades, os seguintes vetores especiais foram criados. Em cada caso, após os DNA terem sido introduzidos nas bactérias, eles se replicam como plasmídeos grandes.

Cosmídeos são vetores que levam insertos de 35 a 45 kb. Eles são híbridos criados de DNA de fago λ e DNA plasmidial bacteriano. Os cosmídeos são inseridos em partículas de fago λ , que atuam como “seringas” que introduzem esses grandes pedaços de DNA recombinante em *E. coli* receptoras. O plasmídeo componente do cosmídeo fornece seqüências necessárias para a replicação deste. Após estarem na célula, esses híbridos formam moléculas que se replicam extracromossomicamente do mesmo modo que os plasmídeos. Os vetores **PAC (cromossomo artificial baseado em P1)** levam o DNA por um sistema similar, mas podem aceitar insertos que variam de 80 a 100 kb. Nesse caso, o vetor é um derivado do bacteriófago P1, um tipo que naturalmente tem um genoma maior do que o fago λ . Os vetores **BAC (cromossomo artificial bacteriano)**, derivados do plasmídeo F, podem levar insertos que variam de 150 a 300 kb (Figura 20.7). O DNA a ser clonado é inserido no plasmídeo, e esse grande DNA recombinante circular é introduzido na bactéria por um tipo especial de transformação. Os BAC são os vetores “burros de carga” para a extensa clonagem necessária em projetos de seqüenciamento genômico em grande escala (discutidos no Capítulo 13).

Finalmente, os insertos maiores que 300 kb precisam de um sistema vetor eucariótico chamado de **YAC (cromossomo artificial de levedura)**. Para construir um YAC, um centrômero cromossômico de levedura e origens de replicação são adicionados a um plasmídeo bacteriano. O plasmídeo é mudado da forma circular para a linear, e o DNA dos telômeros de levedura é adicionado. Essa construção comporta-se de muitos modos como um pequeno cromossomo de levedura na mitose e meiose.

Para clonar o gene de insulina humana, foi selecionado um plasmídeo hospedeiro para levar os insertos de cDNA relativamente curtos com aproximadamente 450 pb. Esse hospedeiro é um tipo especial de plasmídeo chamado de *vetor de expressão*. Os vetores de expressão contêm promotores bacterianos que iniciarão a transcrição em níveis altos, quando o regulador alostérico apropriado for adicionado ao meio de cultura. O vetor de expressão induz cada bactéria

Clonagem em fago λ

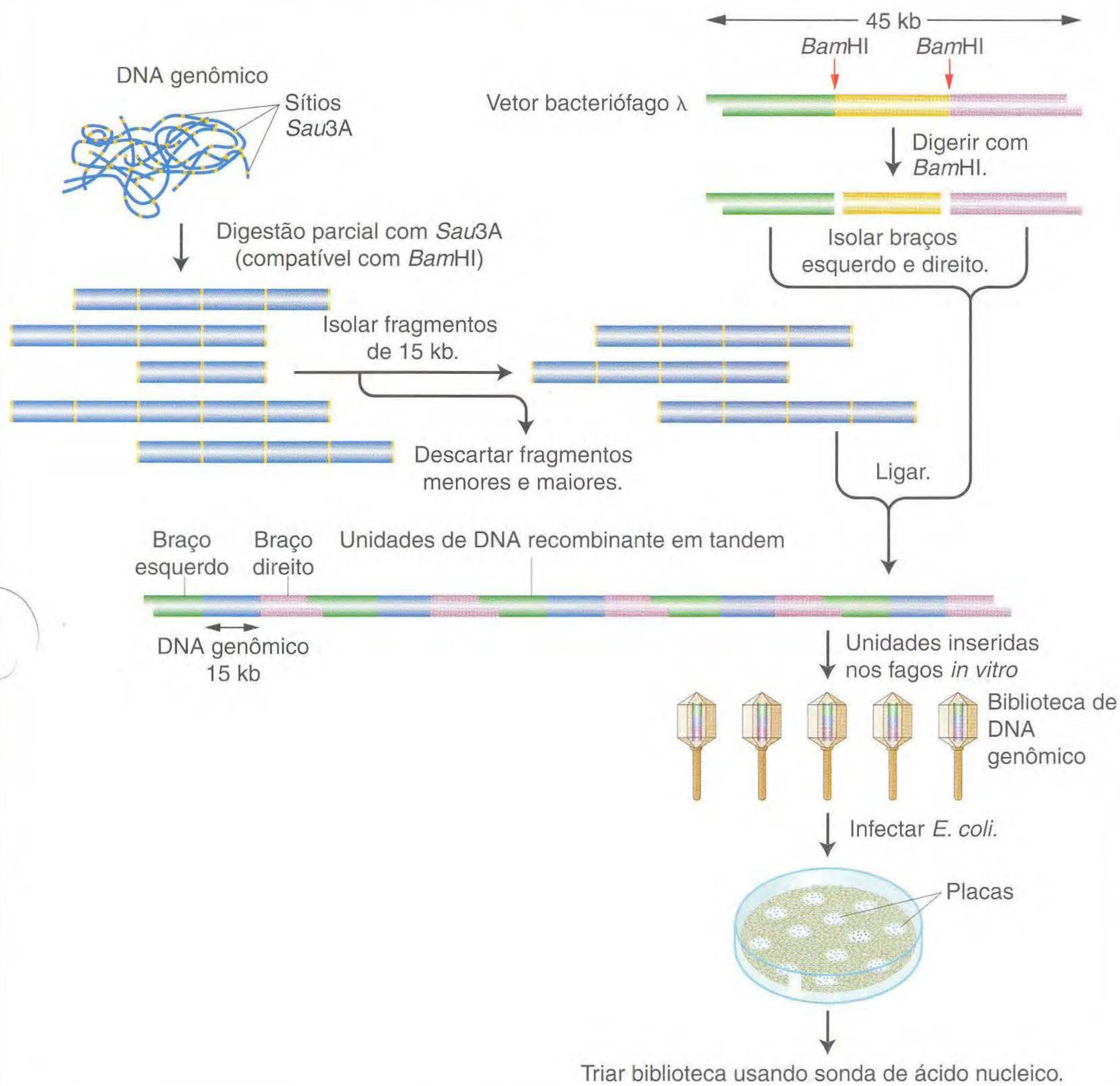


FIG. 20.6 Para clonar em fago λ , uma região central não-essencial do cromossomo do fago é removida e as pontas são ligadas a fragmentos aleatórios de 15 kb do DNA doador. Um multímero linear (concatenado) é formado, o qual então é inserido nas cabeças dos fagos, um monômero de cada vez, usando um sistema de embalagem *in vitro*. [De J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller, *Recombinant DNA*, 2nd ed. Copyright 1992 by Scientific American Books.]

contendo plasmídeo a produzir grandes quantidades de insulina humana recombinante.

Entrada de moléculas recombinantes nas bactérias

As moléculas exógenas de DNA podem entrar em uma bactéria por duas vias básicas: transformação e transdução (Figura 20.8). Na transformação, as bactérias são banhadas em uma solução contendo a molécula de DNA recombinante, que entra na célula e forma um cromossomo plasmídeo (Figura 20.8a). Na transdução, a molécula recombinante é combinada com a cabeça do fago e com proteínas da cauda. Esses fagos construídos são então misturados com as bactérias, injetando sua carga de DNA nas células bacterianas. Se o resultado da injeção é a introdução de um novo plasmídeo recombinante (Figura 20.8b) ou a produção de fagos de prole levando a molécula de DNA recombinante (Figura 20.8c), isso depende do vetor usado. Se for este último, os fagos livres resultantes então infectam a bactéria vizinha. Quando é usado o fago

λ , por repetidas rodadas de reinfecção, forma-se uma placa cheia de partículas de fago de cada bactéria inicial que foi infectada. Cada partícula de fago em uma placa contém uma cópia do cromossomo λ recombinante original.

Recuperação de moléculas recombinantes amplificadas

O DNA recombinante embalado em partículas de fago é facilmente obtido conectando um lisado de fago e isolando o DNA que eles contêm. Para plasmídeos, as bactérias são química ou mecanicamente rompidas. O DNA plasmidial recombinante é separado do cromossomo bacteriano muito maior por centrifugação, eletroforese ou outras técnicas seletivas.

A clonagem gênica é feita pela introdução de vetores recombinantes únicos em bactérias receptoras, seguida de amplificação dessas moléculas como um resultado da tendência natural desses vetores em se replicar.

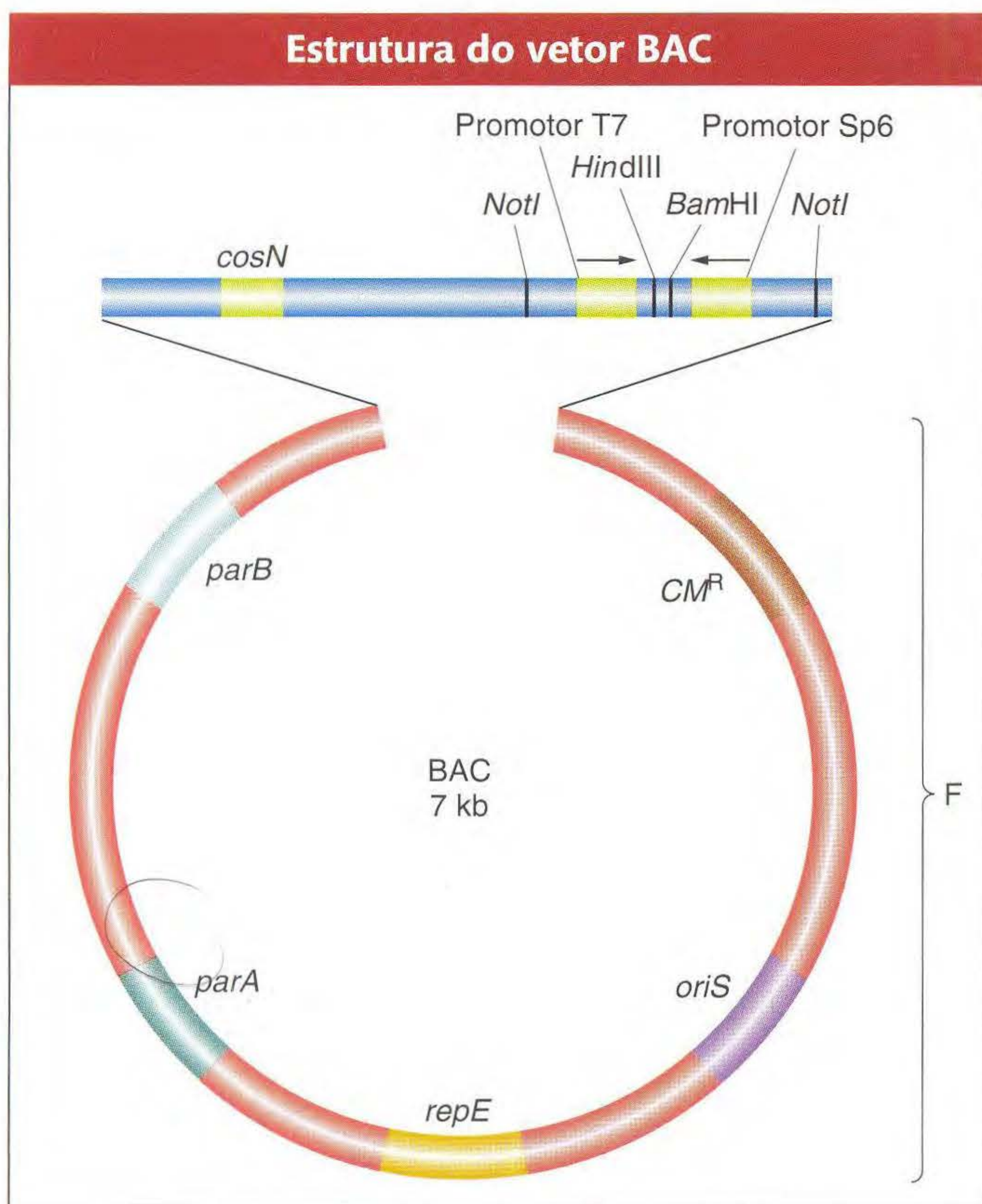


FIG. 20.7 Um cromossomo artificial bacteriano (BAC) é usado para clonar grandes fragmentos de DNA doador. CM^R é um marcador selecionável para resistência ao cloranfenicol. *oriS*, *repE*, *parA* e *parB* são genes de plasmídeo F para replicação e regulação do número de cópias, *cosN* é o sítio *cos* do fago λ . *HindIII* e *BamHI* são sítios de clonagem nos quais o DNA doador é inserido. Os dois promotores são para transcrever o fragmento inserido. Os sítios *NotI* são usados para remover o fragmento inserido. F = plasmídeo F.

Construção de bibliotecas genômicas e de cDNA

Já vimos como fazer e amplificar moléculas de DNA recombinante individuais. Qualquer clone representa uma pequena

parte do genoma de um organismo ou apenas uma das milhares de moléculas de mRNA que o organismo pode sintetizar. Para garantir que clonamos o segmento de DNA de interesse, temos que fazer coleções de segmentos de DNA que sejam inclusivas. Por exemplo, pegamos todo o DNA de um genoma e o quebramos em segmentos do tamanho certo para nosso vetor de clonagem; a seguir, inserimos cada segmento em uma cópia diferente do vetor, criando assim uma coleção de moléculas de DNA recombinante que, em conjunto, representam todo o genoma. Nós então transformamos ou transduzimos essas moléculas em células bacterianas receptoras separadas, onde elas são amplificadas. A coleção resultante de bactérias ou bacteriófagos portadores de DNA é chamada de **biblioteca genômica**. Se estivermos usando um vetor de clonagem que aceita um inserto de tamanho médio de 10 kb, e se todo o genoma tem 100.000 kb de tamanho (o tamanho aproximado do genoma do nematódeo *Caenorhabditis elegans*), então 10.000 clones recombinantes independentes representam um genoma de DNA. Para garantir que todas as seqüências do genoma que pode ser clonado estão contidas dentro de uma coleção, as bibliotecas genômicas representam tipicamente um segmento médio do genoma de pelo menos cinco vezes (assim, em nosso exemplo, existirão 50.000 clones independentes na biblioteca genômica). Essa representação repetida torna altamente improvável que, por acaso, uma seqüência não seja representada pelo menos uma vez na biblioteca.

Similarmente, as coleções representativas de insertos de cDNA requerem dezenas, centenas ou milhares de clones independentes de cDNA. Essas coleções são **bibliotecas de cDNA** e representam apenas as regiões codificantes de proteína do genoma. Uma biblioteca razoável de cDNA é baseada em amostras de mRNA de tecidos diferentes, estágios desenvolvimentais diferentes e organismos cultivados em condições ambientais diferentes.

A escolha de construir uma biblioteca de DNA genômico ou uma biblioteca de cDNA depende da situação. Se estamos procurando um gene específico que é ativo em um tipo específico de tecido em uma planta ou animal, então faz

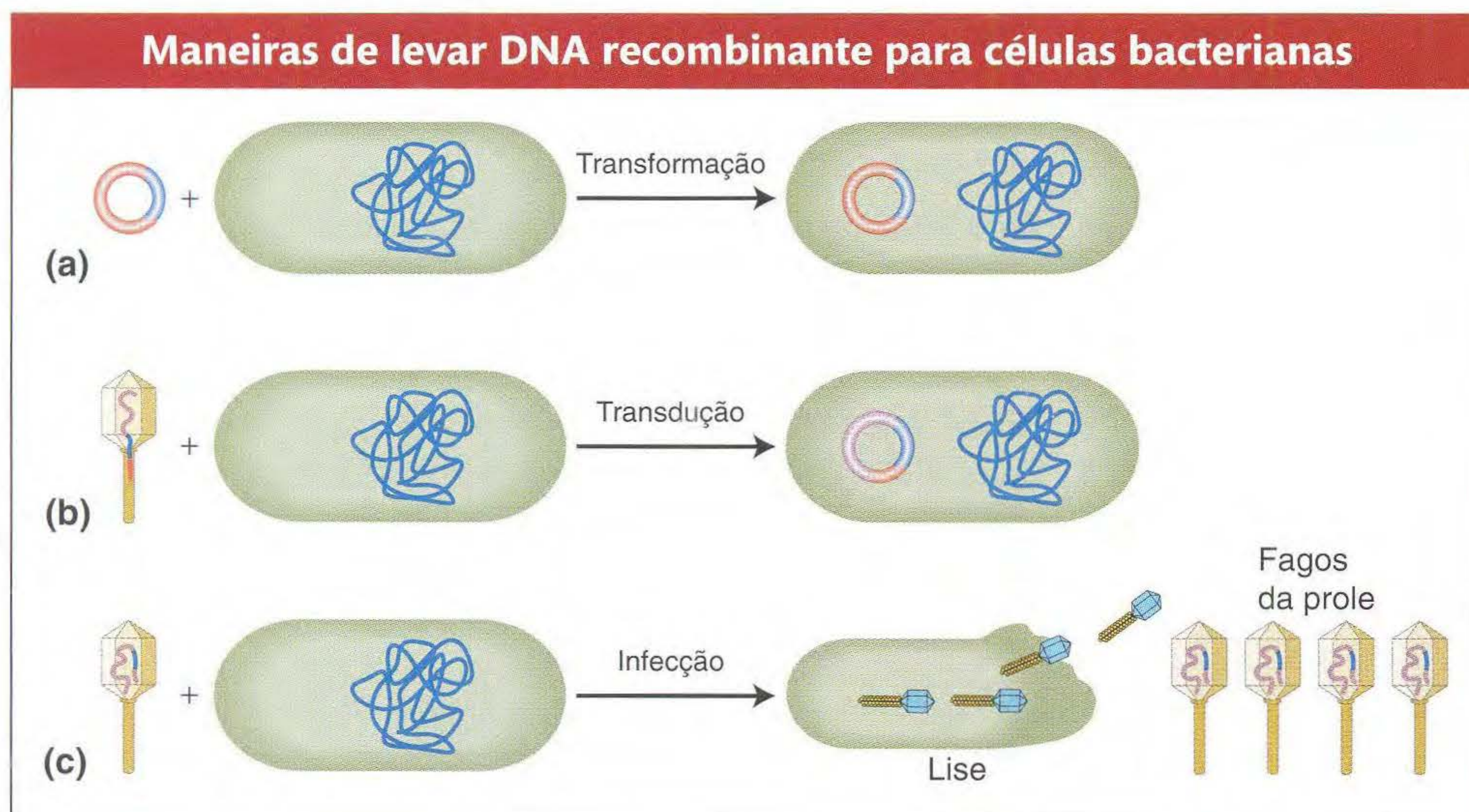


FIG. 20.8 O DNA recombinante pode ser levado para bactérias por transformação, transdução ou infecção com um fago. (a) Um plasmídeo vetor é levado por transformação mediada por DNA. (b) Alguns vetores, tais como cosmídeos, são levados dentro das cabeças dos bacteriófagos (transdução); entretanto, após ter sido injetado na bactéria, eles formam círculos e se replicam como grandes plasmídeos. (c) Os vetores bacteriófagos, tais como fago λ , infectam e lisam a bactéria, liberando um clone de fagos da prole, todos levando moléculas de DNA recombinante idênticas dentro do genoma do fago.

sentido construir uma biblioteca de cDNA de uma amostra desse tecido. Por exemplo, suponha que queremos identificar cDNA correspondentes a mRNA de insulina. As células das ilhotas B do pâncreas são a fonte mais abundante de insulina, e, assim, os mRNA das células pancreáticas são a fonte apropriada para uma biblioteca de cDNA porque esses mRNA devem ser enriquecidos pelo gene em questão. Uma biblioteca de cDNA representa um subgrupo de regiões transcritas do genoma; logo, será inevitavelmente menor do que uma biblioteca genômica completa. Embora as bibliotecas genômicas sejam maiores, elas têm o benefício de conter genes em sua forma nativa, incluindo íntrons e seqüências reguladoras não-transcritas. Uma biblioteca genômica é necessária em algum estágio como um prelúdio para clonar um gene inteiro ou um genoma inteiro.

A tarefa de isolar um clone de um gene específico começa fazendo-se uma biblioteca de DNA genômico ou cDNA, se possível enriquecida de seqüências contendo o gene em questão.

Encontrar um clone específico de interesse

A produção de uma biblioteca como descrita às vezes é chamada de clonagem “shotgun”, porque o experimentador clona uma grande amostra de fragmentos e espera que um dos clones contenha um “hit”, o gene desejado. A tarefa então é encontrar esse clone particular, considerada a seguir.

Encontrar clones específicos usando sondas Uma biblioteca pode conter até centenas ou milhares de fragmentos clonados. Essa enorme coleção de fragmentos deve ser triada para encontrar a molécula de DNA recombinante contendo o gene de interesse do pesquisador. Tal triagem é feita usando uma **sonda** específica que encontrará e marcará apenas o clone desejado. Existem dois tipos de sondas: (1) as que reconhecem uma seqüência específica de ácido nucleico e (2) as que reconhecem uma proteína específica.

Sondas para encontrar DNA As sondas para DNA fazem uso do poder de complementariedade de bases. Dois ácidos nucleicos unifilamentares com complementariedade total ou parcial de seqüência de bases se “encontrarão” na solução por colisão aleatória. Após serem unidos, o híbrido bifilamentar assim formado é estável. Esse enfoque é um meio poderoso de encontrar seqüências específicas de interesse. As sondas de DNA requerem que todas as moléculas sejam tornadas unifilamentares por aquecimento. Uma sonda unifilamentar, marcada radioativa ou quimicamente, é enviada para encontrar sua seqüência-alvo complementar em uma população de DNA tal como uma biblioteca. As sondas tão pequenas quanto 15 a 20 pares de bases irão hibridizar-se a seqüências complementares dentro de clones maiores de DNA. Assim, as sondas podem ser vistas como “iscas” para identificar “presas” muito maiores.

A identificação de um clone específico em uma biblioteca é um processo de duas etapas (Figura 20.9). Primeiro, as

colônias ou placas da biblioteca em uma placa de Petri são transferidas para uma membrana absorvente simplesmente colocando a membrana na superfície do meio. A membrana é removida, as colônias ou placas grudadas na superfície são lisadas *in situ*, e o DNA é desnaturado. Segundo, a membrana é banhada com uma solução de sonda unifilamentar específica para o DNA que está sendo procurado. Geralmente, a sonda em si é um trecho de DNA clonado que tem uma seqüência homóloga à do gene desejado. A sonda deve ser marcada com um isótopo radioativo ou um corante fluorescente. Assim, a posição de um clone positivo ficará clara pela posição da marcação concentrada radioativa ou fluorescente. Para marcadores radioativos, a membrana é colocada em um filme de raios X, e o decaimento do radioisótopo produz partículas subatômicas que “expõem” o filme, produzindo neste um ponto escuro adjacente ao local da concentração de radioisótopos. Tal filme exposto é chamado de **auto-radiograma**. Se for usado um corante fluorescente como marcador, a membrana é exposta ao comprimento de onda correto de luz para ativar a fluorescência do corante, e é tirada uma foto da membrana para registrar o local do corante fluorescente.

De onde veio o DNA para fazer a sonda? O DNA pode vir de várias fontes.

- *DNA complementar de um tecido que expressa um gene de interesse em alto nível.* Para o gene de insulina, o pâncreas seria a escolha óbvia.
- *Um gene homólogo de um organismo correlato.* Esse método depende da conservação evolutiva das seqüências de DNA ao longo do tempo. Muito embora o DNA da sonda e o DNA do clone desejado possam não ser idênticos, em geral eles são suficientemente similares para promover hibridização. O método é humoristicamente chamado de “clone por telefone” porque, se você puder telefonar para um colega que tenha o clone de seu gene de um organismo relacionado, sua tarefa de clonagem é tornada relativamente fácil.
- *O produto proteico do gene de interesse.* A seqüência de aminoácidos de parte da proteína é retrotraduzida, usando o quadro do código genético ao reverso (do aminoácido para o códon), para obter a seqüência de DNA nele codificada. É então criada uma sonda sintética de DNA que corresponde a essa seqüência. Lembre, entretanto, que o código genético é redundante, isto é, esses aminoácidos são codificados por múltiplos códons. Assim, várias seqüências possíveis de DNA podem teoricamente codificar a proteína em questão, mas apenas uma dessas seqüências de DNA está presente no gene que de fato codifica a proteína. Para contornar esse problema, é selecionado um curto trecho de aminoácidos com a mínima redundância. É então criado um conjunto misto de sondas contendo todas as seqüências de DNA que possam codificar essa seqüência de aminoácidos. Esse “coquetel” de oligonucleotídeos é usado como uma sonda. O filamento correto dentro desse coquetel encontra o gene de interesse. Cerca de 20 nucleotídeos englobam especificidade suficiente para se hibridizar a uma única seqüência complementar de DNA na biblioteca.

Encontrar o clone de interesse usando sondas de DNA ou RNA

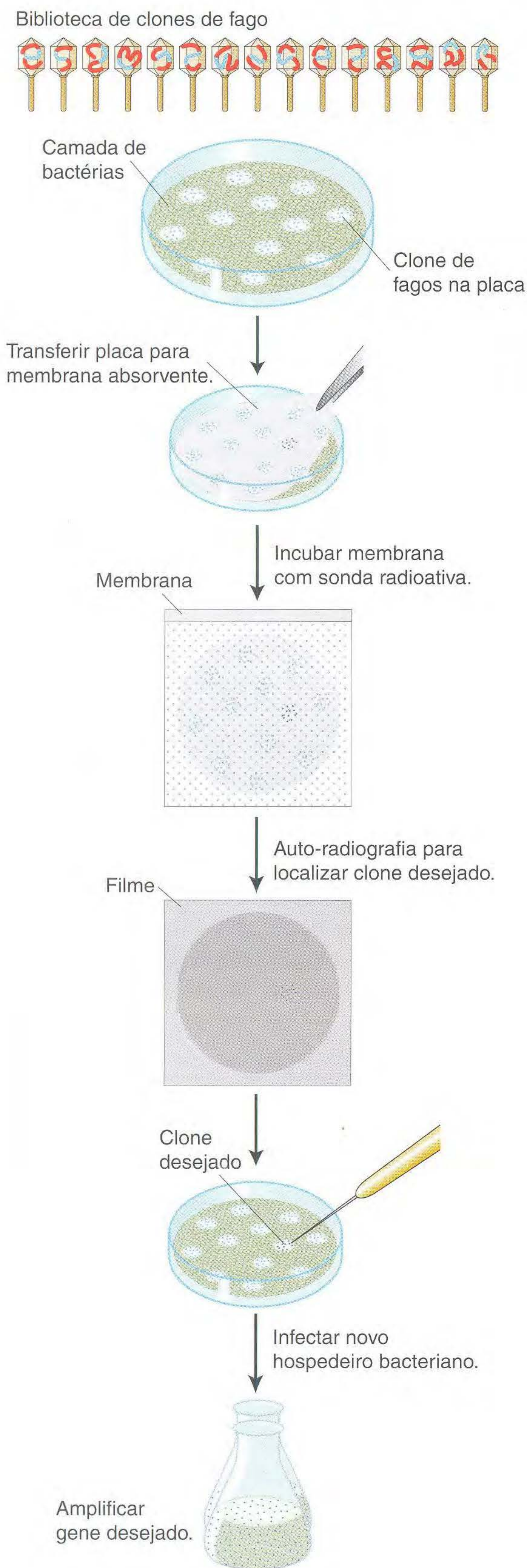


FIG. 20.9 O clone levando um gene de interesse é identificado sondando uma biblioteca genômica, nesse caso feita por genes clonados nos bacteriófagos λ , com DNA ou RNA conhecido como relacionado ao gene desejado. Uma sonda radioativa hibridiza-se com qualquer DNA recombinante incorporando uma sequência de DNA compatível, e a posição do clone tendo o DNA é revelada por auto-radiografia. Agora o clone desejado pode ser selecionado pelo ponto correspondente na placa de Petri e transferido para um novo hospedeiro bacteriano, de modo que um gene puro pode ser manufaturado. [De R.A. Weinberg, "A Molecular Basis of Cancer," and P. Leder, "The Genetics of Antibody Diversity." Copyright 1983, 1982 by Scientific American, Inc. Todos os direitos reservados.]

- **RNA livre marcado.** Esse tipo de sonda é possível apenas quando uma população quase pura de moléculas idênticas de RNA pode ser isolada, tal como rRNA.

Sondas para encontrar proteínas Se o produto proteico de um gene é conhecido e isolado sob forma pura, então essa proteína pode ser usada para detectar o clone do gene correspondente em uma biblioteca. O processo, descrito na Figura 20.10, requer dois componentes. Primeiro, requer uma biblioteca de expressão, feita usando vetores de expressão. Para fazer a biblioteca, o cDNA é inserido no vetor na trinca correta da matriz de leitura com uma proteína bacteriana (nesse caso, β -galactosidase), e as células contendo o vetor e seu inserto produzem uma proteína de "fusão" que é, em parte, uma tradução da inserção de cDNA e, em parte, um trecho da β -galactosidase normal. Segundo, o processo requer um **anticorpo** ao produto proteico específico do gene de interesse. (Um anticorpo é uma proteína feita pelo sistema imunológico do animal que se liga com alta afinidade a determinada molécula.) O anticorpo é usado para triar a biblioteca de expressão dessa proteína. Uma membrana é colocada sobre a superfície do meio e removida de modo que algumas células de cada colônia agora estão ligadas à membrana em locais que correspondem a suas posições na placa de Petri original (veja Figura 20.10). A membrana "impressa" é então seca e banhada em uma solução do anticorpo, que se ligará à impressão de qualquer colônia que contenha a proteína de fusão de interesse. Os clones positivos são revelados por um anticorpo secundário marcado que se liga ao primeiro anticorpo. Detectando a proteína correta, o anticorpo efetivamente identifica o clone contendo o gene que deve ter produzido essa proteína e, portanto, contém o cDNA desejado.

Um gene clonado pode ser selecionado de uma biblioteca usando-se sondas para a sequência de DNA do gene ou para o produto proteico do gene.

Sonda para encontrar um ácido nucleico específico em uma mistura Como veremos mais adiante neste capítulo, durante a manipulação do gene e do genoma, geralmente é necessário detectar e isolar uma molécula específica de DNA ou RNA em meio a uma mistura complexa.

O método mais amplamente usado para detectar uma molécula dentro de uma mistura é o *blotting* ("borrão"), que

Encontrar o clone de interesse usando anticorpos

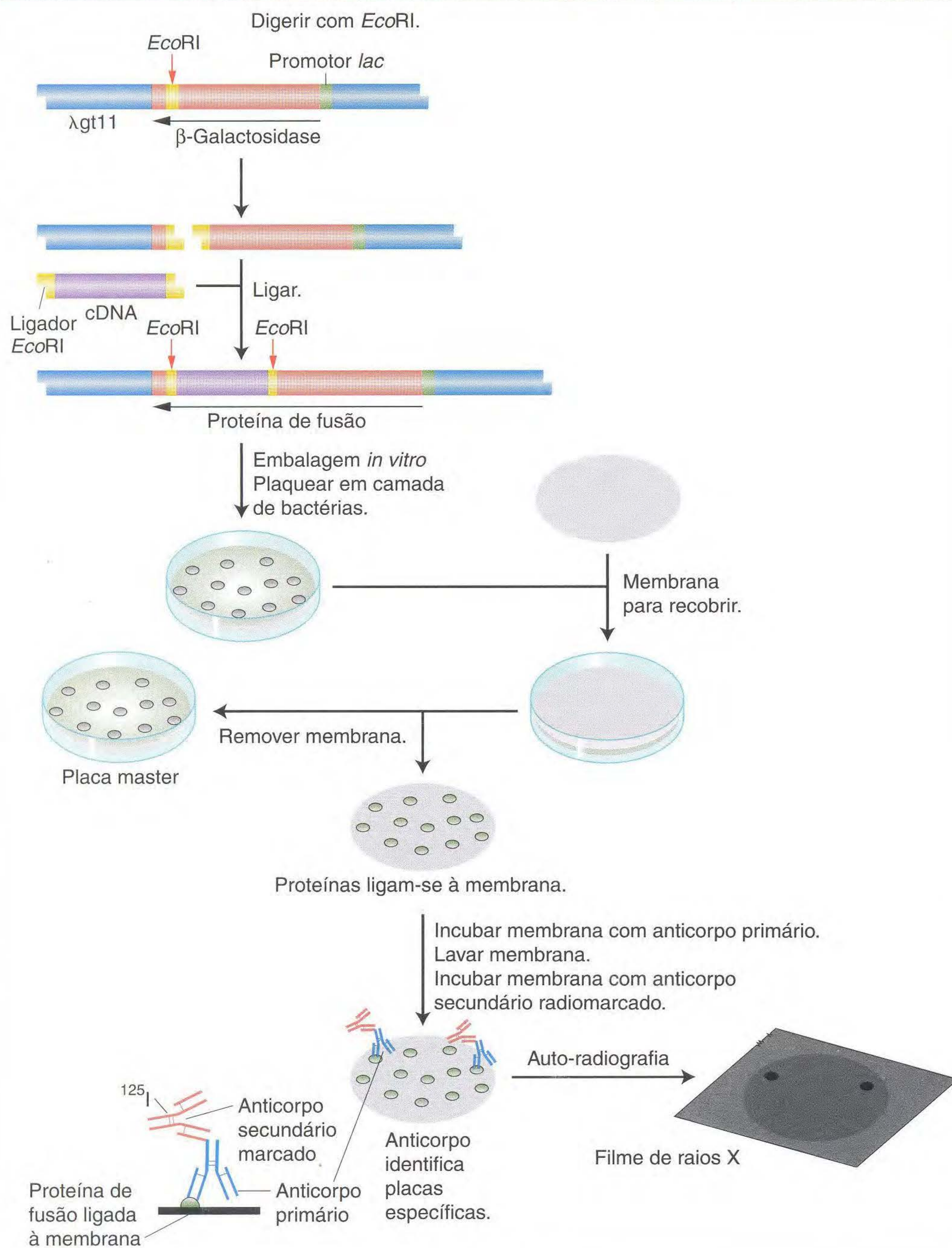


FIG. 20.10 Para encontrar o clone de interesse, uma biblioteca de expressão feita com um vetor de fago λ especial chamado de λ gt11 é triada com um anticorpo específico de proteína. Após os anticorpos não-ligados terem sido removidos da membrana, os anticorpos ligados são visualizados pela ligação de um anticorpo secundário radioativo. [De J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller, *Recombinant DNA*, 2nd ed. Copyright 1992 by Scientific American Books.]

começa com a **eletroforese em gel** para separar as moléculas da mistura. Vejamos o DNA primeiro. Uma mistura de moléculas lineares de DNA é colocada em uma cuba de um gel de agarose, e a cuba é ligada ao catódio de um campo elétrico. Como as moléculas de DNA contêm cargas, os fragmentos irão migrar pelo gel para o anódio em velocidades inversamente dependentes de seu tamanho (Figura 20.11). Portanto, os fragmentos em classes de tamanhos distintos formarão bandas no gel. As bandas podem ser visualizadas corando-se o DNA com brometo de etídio, que faz com que o DNA fluoresça na luz ultravioleta. O tamanho absoluto de

cada fragmento na mistura pode ser determinado comparando sua distância de migração com um conjunto de fragmentos padrão de tamanhos conhecidos. Se as bandas são bem separadas, uma banda individual pode ser cortada do gel, e a amostra de DNA pode ser purificada da matriz do gel. Portanto, a eletroforese de DNA pode ser diagnóstica (mostrando tamanhos e quantidades relativas de fragmentos de DNA presentes) ou preparativa (útil em isolar fragmentos específicos de DNA).

O DNA genômico digerido por enzimas de restrição em geral produz tantos fragmentos que a eletroforese produz um

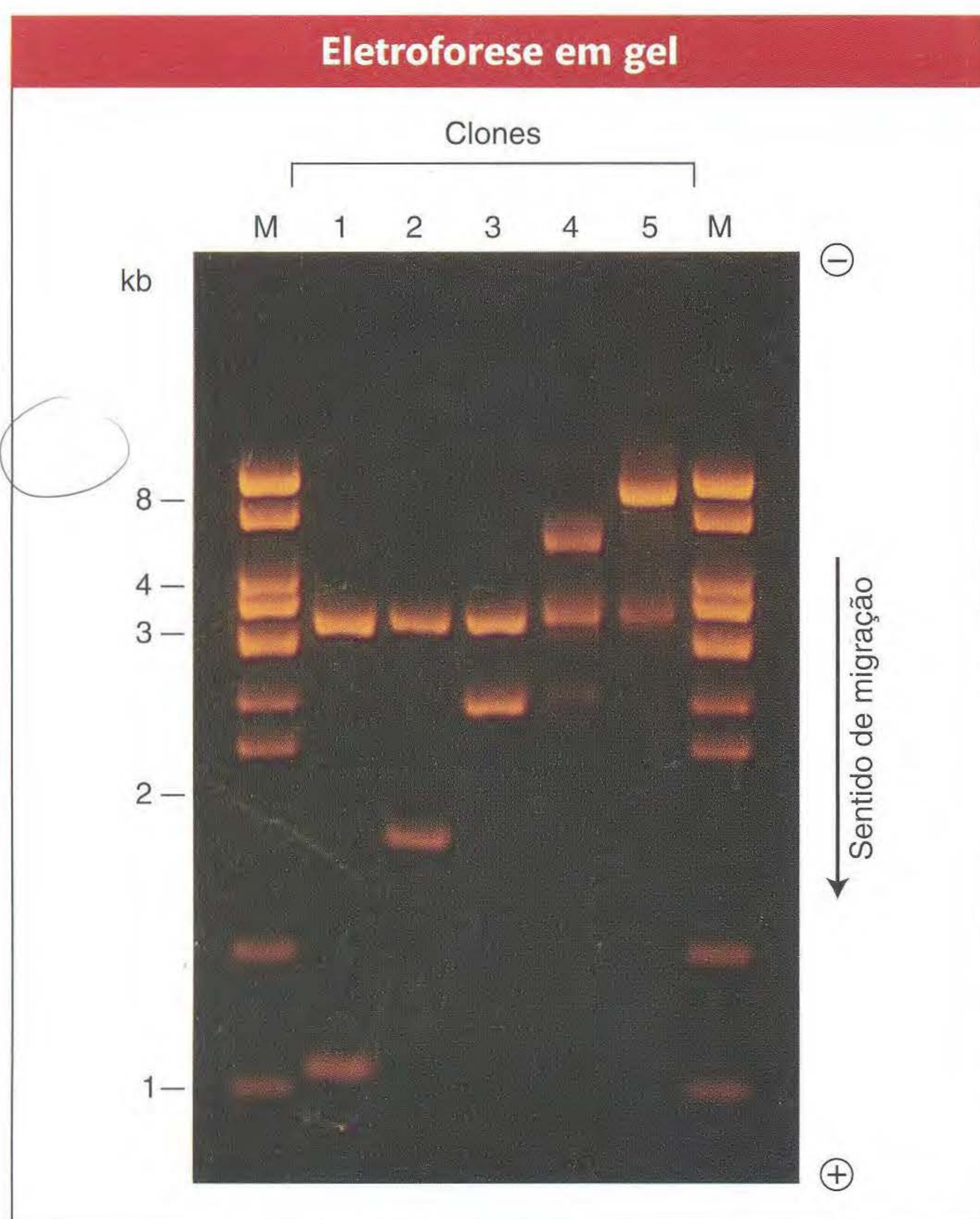


FIG. 20.11 As misturas de fragmentos de DNA de vários tamanhos foram separadas eletroforéticamente em gel de agarose. As amostras são cinco vetores recombinantes tratados com *EcoRI*. As misturas são aplicadas a cubas no topo do gel, e os fragmentos movem-se sob a influência de um campo isoeletrico para posições diferentes dependentes do tamanho (e, portanto, do número de cargas). As bandas de DNA foram visualizadas corando com brometo de etídio e fotografando sob luz ultravioleta. (A letra "M" representa colunas contendo fragmentos padrão que agem como marcadores para avaliar o tamanho do DNA.) [De H. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, and J. Darnell, *Molecular Cell Biology*, 3rd ed. Copyright 1995 by Scientific American Books.]

esfregaço contínuo de DNA sem bandas distintas. Uma sonda pode identificar um fragmento nessa mistura usando-se uma técnica desenvolvida por E.M. Southern chamada de **transferência de Southern** (Figura 20.12). Como a identificação de clones (veja Figura 20.9), essa técnica gera um *imprint* de moléculas de DNA em uma membrana usando a membrana para "transferir" o gel após a eletroforese estar completa. O DNA deve ser primeiro desnaturado, o que lhe permite grudar na membrana. Então a membrana é hibridizada com a sonda marcada. Um auto-radiograma ou uma foto de bandas fluorescentes revelará a presença de quaisquer bandas no gel que sejam complementares à sonda. Se apropriado, essas bandas podem ser removidas do gel e processadas.

A técnica de transferência de Southern pode ser ampliada para detectar uma molécula específica de RNA de uma mistura de RNA fracionados em um gel. Essa técnica é chamada de **transferência Northern** (graças ao senso de humor de alguns cientistas) para contrastar com a técnica de trans-

ferência de Southern usada para análise de DNA. O RNA separado por eletroforese é transferido para uma membrana e sondado do mesmo modo que o DNA na transferência de Southern. Uma aplicação da análise Northern é determinar se um gene específico é transcrito em um certo tecido ou sob certas condições ambientais.

Assim, vemos que o DNA clonado tem amplas aplicações como uma sonda usada para detectar um clone específico, um fragmento de DNA ou uma molécula de RNA. Em todos esses casos, note que a técnica novamente explora a habilidade dos ácidos nucleicos com sequências *complementares* em encontrar e se ligar uma à outra.

As técnicas de DNA recombinante que dependem de complementariedade a uma sonda clonada de DNA incluem os sistemas de *blotting* e hibridização para a identificação de clones específicos, fragmentos de restrição ou mRNA, ou para medida do tamanho de DNA ou RNA específicos.

Encontrando clones específicos por complementação funcional Em muitos casos, não temos uma sonda para começar, mas temos uma mutação recessiva no gene de interesse. Se formos capazes de introduzir um DNA funcional na espécie portadora desse alelo (veja Seção 20.6, Engenharia Genética), podemos detectar clones específicos em uma biblioteca de bactérias ou fagos por sua habilidade em restaurar a função eliminada pela mutação recessiva nesse organismo. Esse procedimento é chamado de **complementação funcional** ou **recuperação de mutante**. O esquema geral do procedimento é o seguinte:

Fazer uma biblioteca de bactérias ou fagos contendo insertos de DNA doador recombinante tipo selvagem a^+ .



Transformar células de uma linhagem celular mutante recessiva a^- usando o DNA de clones individuais na biblioteca.



Identificar clones da biblioteca que produzem células transformadas com o fenótipo dominante a^+ .



Recuperar o gene a^+ do clone bem-sucedido de fago ou bactéria.

Encontrando clones específicos com base na localização do mapa genético – clonagem posicional A informação sobre a posição de um gene no genoma pode ser usada para evitar o duro trabalho de triar uma biblioteca inteira para encontrar o clone de interesse. **Clonagem posicional** é um termo que pode ser aplicado a qualquer método para encontrar um clone específico que faz uso da informação sobre a posição do gene em seu cromossomo. Dois elementos são necessários para a clonagem posicional:

- *Alguns marcoss genéticos podem delimitar onde o gene pode estar.* Se possível, os marcoss em ambos os lados do gene

Encontrar ácidos nucleicos específicos usando eletroforese em gel e transferência

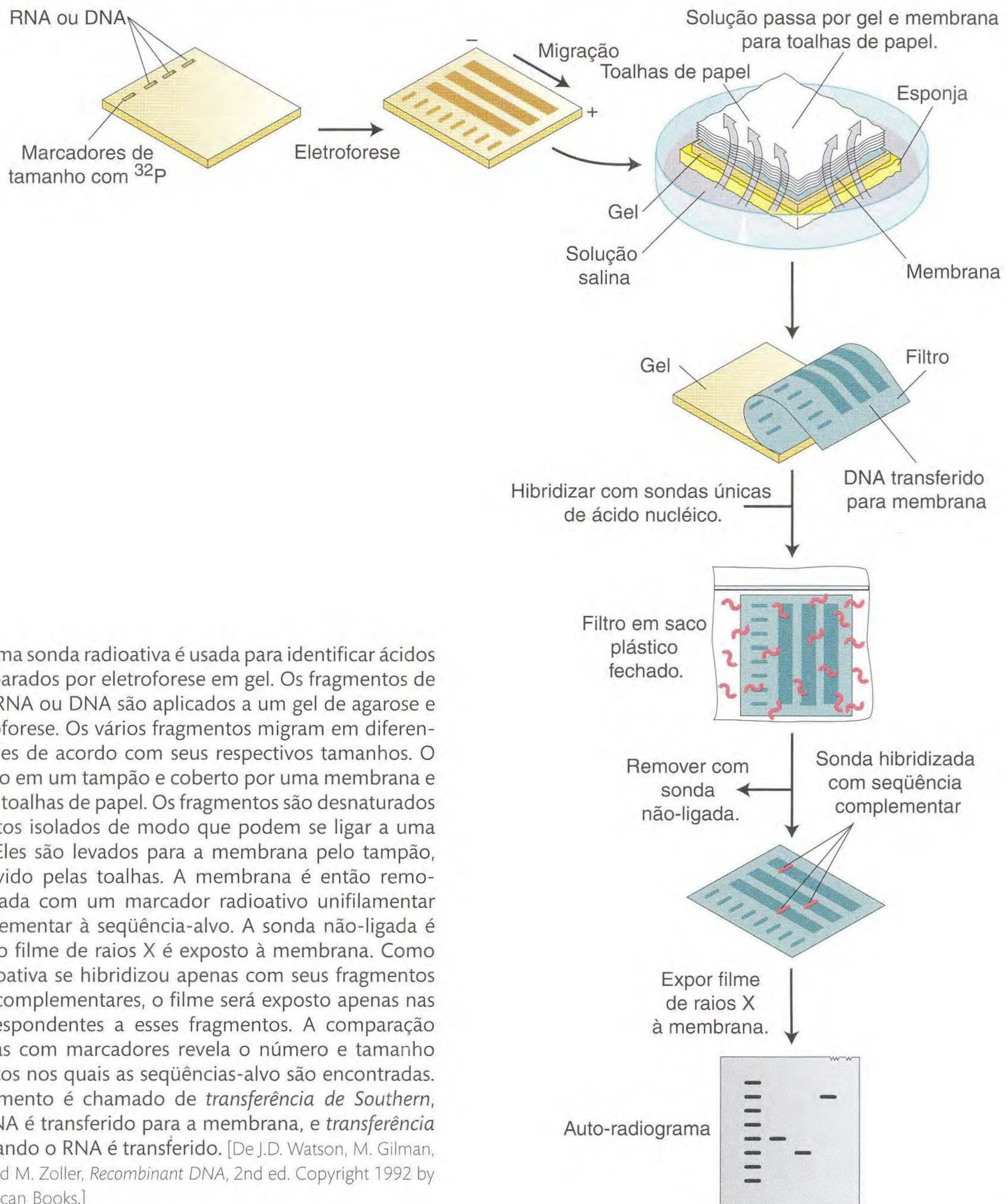


FIG. 20.12 Uma sonda radioativa é usada para identificar ácidos nucleicos separados por eletroforese em gel. Os fragmentos de restrição de RNA ou DNA são aplicados a um gel de agarose e sofrem eletroforese. Os vários fragmentos migram em diferentes velocidades de acordo com seus respectivos tamanhos. O gel é colocado em um tampão e coberto por uma membrana e uma pilha de toalhas de papel. Os fragmentos são desnaturados para filamentos isolados de modo que podem se ligar a uma membrana. Eles são levados para a membrana pelo tampão, que é absorvido pelas toalhas. A membrana é então removida e incubada com um marcador radioativo unifilamentar que é complementar à sequência-alvo. A sonda não-ligada é removida, e o filme de raios X é exposto à membrana. Como a sonda radioativa se hibridizou apenas com seus fragmentos de restrição complementares, o filme será exposto apenas nas bandas correspondentes a esses fragmentos. A comparação dessas bandas com marcadores revela o número e tamanho dos fragmentos nos quais as sequências-alvo são encontradas. Esse procedimento é chamado de *transferência de Southern*, quando o DNA é transferido para a membrana, e *transferência Northern*, quando o RNA é transferido. [De J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller, *Recombinant DNA*, 2nd ed. Copyright 1992 by Scientific American Books.]

de interesse são melhores, pois delimitam a possível localização desse gene. Os marcadores podem ser RFLP (polimorfismos de comprimento de fragmento de restrição), SNP (polimorfismos de um só nucleotídeo) ou outros polimorfismos moleculares (veja Capítulos 4 e 13), ou podem ser pontos de quebra cromossômica bem mapeados (veja Capítulo 16).

- A habilidade em investigar o segmento contínuo de DNA que se situa entre os marcadores genéticos delimitantes. Em organismos-modelo, os genes nesse bloco de DNA são conhecidos pela sequência do genoma (veja Capítulo 13).

A partir desses genes, podem ser escolhidos candidatos que podem representar o gene que está sendo procurado. Para outras espécies, é usado um procedimento chamado de **andar no cromossomo** para encontrar e ordenar os clones que caem entre marcadores genéticos. A Figura 20.13 resume o procedimento. A ideia básica é usar a sequência do marco próximo como uma sonda para identificar um segundo grupo de clones que se superpõem ao clone marcador contendo o marco mas que se prolongam além dele em uma das duas direções (no sentido do marco ou se afastando do alvo). Os fragmentos termi-

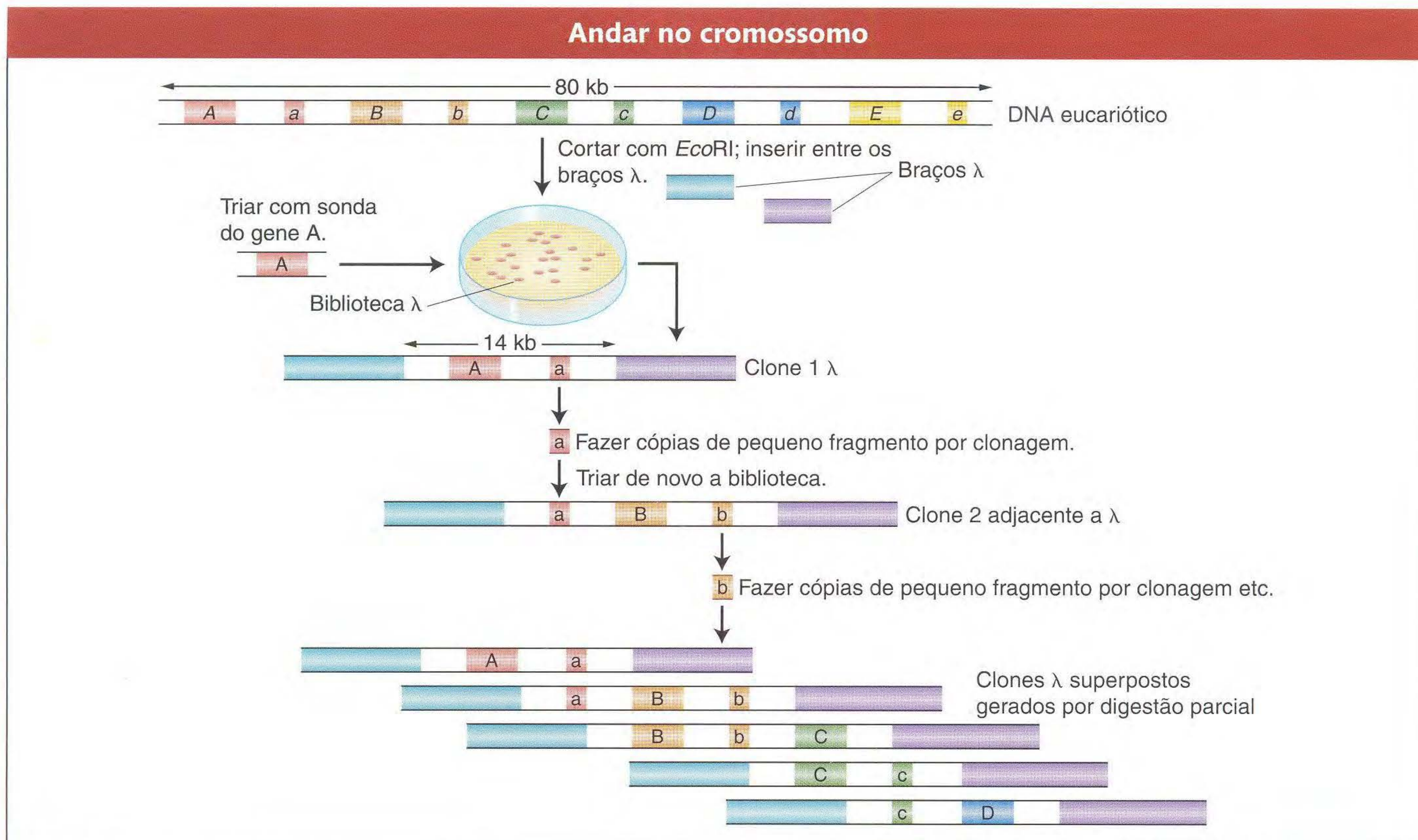


FIG. 20.13 Essa caminhada no cromossomo começa com um fago recombinante obtido de uma biblioteca de fagos feita de digestão parcial com *EcoRI* de um genoma eucariótico. A fase recombinante é então usada para isolar outro fago recombinante contendo um segmento vizinho de DNA eucariótico. Essa caminhada ilustra como começar em um marcador molecular A e chegar ao gene-alvo D. [De J.D. Watson, J. Tooze, and D.T. Kurtz, *Recombinant DNA: A Short Course*. Copyright 1983 by W.H. Freeman and Company.]

nais do novo conjunto de clones podem ser usados como sondas para identificar um terceiro conjunto de clones superpostos da biblioteca genômica. Desse modo gradativo, um conjunto de clones que representam a região do genoma que se prolonga para fora do clone marcador pode ser triado até que sejam obtidos clones que possam ser mostrados incluindo o gene-alvo, talvez mostrando que ele recupera um mutante do gene-alvo. Esse processo é chamado de “andar” no cromossomo, pois consiste em uma série de passos de um clone adjacente para o seguinte.

► 20.2 Amplificação de DNA *in Vitro*: a Reação em Cadeia da Polimerase

Se soubermos a sequência de pelo menos algumas partes do gene ou sequência de interesse, podemos amplificá-la em um tubo de ensaio em vez de por clonagem. O procedimento é chamado de **reação em cadeia da polimerase (PCR)**. A estratégia básica da PCR é mostrada na Figura 20.14. O processo usa várias cópias de um par de curtos primers sintetizados quimicamente, com 15 a 20 pares de bases de tamanho, cada um ligando-se a uma ponta diferente do gene ou região a ser amplificada. Os dois primers ligam-se a filamentos opo-

tos de DNA, com suas pontas 3' apontando uma para a outra. As polimerases adicionam bases a esses primers e o processo de polimerização vai e volta entre eles, formando um número potencialmente crescente de moléculas bifilamentares de DNA. Os detalhes são os seguintes.

Começamos com uma solução contendo uma fonte de DNA, os primers, os quatro trifosfatos de desoxirribonucleotídeos e uma DNA polimerase especial. O DNA é desnaturado pelo calor, resultando em moléculas unifilamentares de DNA. Os primers hibridizam-se a suas seqüências complementares nas moléculas unifilamentares de DNA em soluções resfriadas. Uma DNA polimerase especial tolerante ao calor replica os segmentos unifilamentares de DNA prolongando-se a partir de um primer. A DNA polimerase *Taq* polimerase, da bactéria *Thermus aquaticus*, é uma dessas enzimas comumente usadas. (Essa bactéria normalmente cresce em fontes térmicas e, assim, desenvolveu proteínas resistentes ao calor extremo. Ela é capaz de sobreviver a altas temperaturas necessárias para desnaturar a dupla hélice de DNA, desnaturando e inativando a DNA polimerase da maioria das espécies.) Novos filamentos complementares são sintetizados como na replicação normal de DNA nas células, formando duas moléculas bifilamentares de DNA idênticas à molécula bifilamentar parental. Essas etapas que levam a uma única replicação do segmento entre os dois primers representam um ciclo.

Reação em cadeia da polimerase

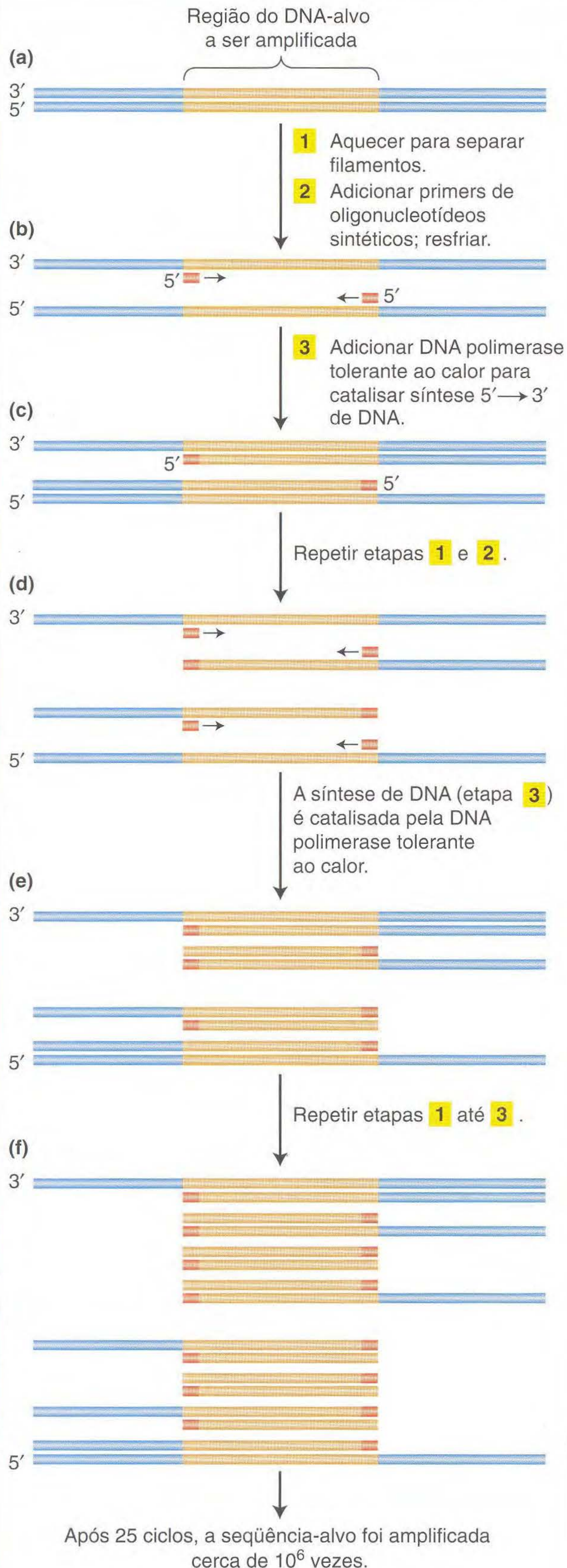


FIG. 20.14 A reação em cadeia da polimerase copia rapidamente uma seqüência de DNA-alvo. (a) O DNA bifilamentar contendo a seqüência-alvo. (b) Dois primers escolhidos ou criados têm seqüências complementares aos sítios de ligação de primer nas pontas 3' do gene-alvo nos dois filamentos. Os filamentos são separados por aquecimento, permitindo que dois primers se helicoidizem aos sítios de ligação de primer. Juntos, os primers flanqueiam a seqüência-alvo. (c) A *Taq* polimerase então sintetiza o primeiro conjunto de filamentos complementares na reação. Esses primeiros dois filamentos são de tamanhos variados, porque não têm um sinal comum de fim. Eles se estendem além das pontas da seqüência-alvo conforme delineado pelos sítios de ligação de primer. (d) Os dois dúplexes são novamente aquecidos, expondo quatro sítios de ligação. Os dois primers novamente se ligam a seus respectivos fragmentos nas pontas 3' da região-alvo. (e) A *Taq* polimerase sintetiza quatro filamentos complementares. Embora os filamentos-molde nesse estágio sejam variáveis em tamanho, dois dos quatro filamentos sintetizados deles são exatamente do tamanho da seqüência-alvo desejada. Esse tamanho preciso é obtido porque cada um desses filamentos começa no sítio de ligação do primer, em uma ponta da seqüência-alvo, e continua até sair do molde, na outra ponta da seqüência. (f) O processo pode ser repetido indefinidamente, cada vez criando mais moléculas de duplo filamento de DNA com a seqüência-alvo. [De D.L. Nelson and M.M. Cox, *Lehninger Principles of Biochemistry*, 4th ed. Copyright 2005 by W.H. Freeman and Company.]

Após a replicação do segmento entre os dois primers estar completa, os dois novos dúplexes são novamente desnaturados por aquecimento para gerar moldes unifilamentares, e ocorre um segundo ciclo de replicação diminuindo a temperatura na presença de todos os componentes necessários para a polimerização. Ciclos repetidos de desnaturação, helicoidização e síntese resultam em um aumento exponencial no número de segmentos replicados. Amplificações de até um milhão de vezes podem ser prontamente obtidas dentro de 1 a 2 horas.

A grande vantagem da PCR é que são necessários menos procedimentos comparados com a clonagem porque a localização dos primers determina a especificidade do segmento de DNA que é amplificado. Se as seqüências que correspondem aos primers estão cada uma presente apenas uma vez no genoma e estão suficientemente próximas (distância máxima de cerca de 2 kb), o *único* segmento de DNA que pode ser amplificado é aquele entre os dois primers. Essa amplificação será obtida mesmo que o segmento de DNA esteja presente em níveis bem baixos (p.ex., uma parte em um milhão) em uma mistura complexa de fragmentos de DNA tal como as preparações do DNA genômico humano.

Como a PCR é uma técnica muito sensível, ela tem muitas aplicações em biologia. Ela pode amplificar seqüências-alvo que estão presentes em números extremamente baixos de cópias em uma amostra, desde que sejam usados primers específicos para essa seqüência rara. Por exemplo, investigadores de crimes podem amplificar segmentos de DNA humano a partir de poucas células foliculares que circundam um único fio de cabelo arrancado.

Embora a sensibilidade e especificidade da PCR sejam vantagens claras, a técnica tem algumas limitações significativas. Para criar os primers de PCR, pelo menos algumas

informações de seqüências devem estar disponíveis sobre o pedaço de DNA que vai ser amplificado. Na ausência de tal informação, a amplificação de PCR não pode ser aplicada.

A reação em cadeia da polimerase usa primers especialmente criados para dirigir a amplificação de regiões específicas de DNA em um tubo de ensaio.

► 20.3 Determinando a Seqüência de Bases de um Segmento de DNA

Após termos clonado nosso gene desejado ou o amplificado usando PCR, começa a tarefa de tentar compreender sua função. A linguagem final do genoma é composta de seqüências de nucleotídeos A, T, C e G. Obter a seqüência completa de nucleotídeos de um segmento de DNA é geralmente uma parte importante da compreensão da organização de um gene e sua regulação, sua relação com outros genes ou a função de seu RNA ou proteína codificada. Na maioria das vezes, traduzir a seqüência do ácido nucleico de uma molécula de cDNA para descobrir a seqüência de aminoácidos de sua cadeia polipeptídica codificada é mais simples que seqüenciar o próprio polipeptídeo. Nesta seção, consideramos as técnicas usadas para ler a seqüência de nucleotídeos do DNA.

Como nas tecnologias do DNA recombinante e PCR, o seqüenciamento do DNA explora a complementariedade de pares de bases juntamente com uma compreensão da bioquímica básica da replicação do DNA. Várias técnicas foram desenvolvidas, mas uma delas é a mais usada. Ela é chamada de **seqüenciamento dideoxi** ou, às vezes, **seqüenciamento de Sanger** em homenagem a seu criador. O termo *dideoxi* vem de um nucleotídeo especial modificado, chamado de trifosfato de dideoxinucleotídeo (genericamente, um ddNTP). Esse nucleotídeo modificado é a chave da técnica de Sanger devido à sua habilidade em bloquear a síntese continuada de DNA. O que é um trifosfato de dideoxinucleotídeo? E como ele bloqueia a síntese de DNA? Um dideoxinucleotídeo não tem o grupo 3'-hidroxila nem o grupo 2'-hidroxila, que também está ausente em desoxinucleotídeo (Figura 20.15). Para ocorrer a síntese de DNA, a DNA polimerase deve catalisar uma reação de condensação entre o grupo 3'-hidroxila do último nucleotídeo adicionado à cadeia crescente e o grupo 5'-fosfato do nucleotídeo seguinte a ser adicionado, liberando água e formando uma ligação fosfodiéster com o átomo 3'-carbono do açúcar adjacente. Como um dideoxinucleotídeo não tem o grupo 3'-hidroxila, essa reação não pode ocorrer, e, portanto, a síntese de DNA é bloqueada no ponto de adição.

A lógica do seqüenciamento dideoxi é direta. Suponha que queremos ler a seqüência de um segmento clonado de DNA de, digamos, 5.000 pares de bases. Pri-

A estrutura de 2',3'-dideoxinucleotídeos

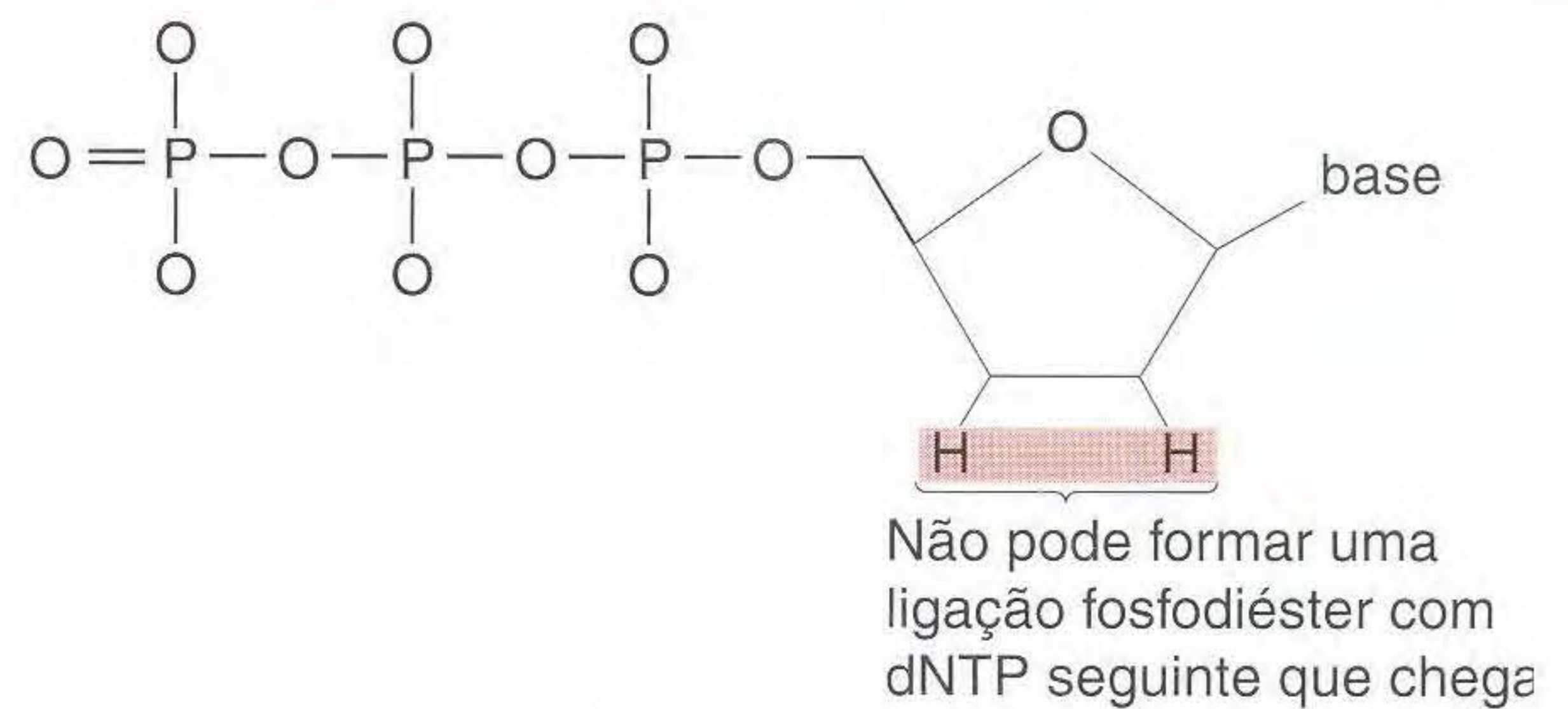


FIG. 20.15 2',3'-Dideoxinucleotídeos, que são empregados no método de seqüenciamento de DNA de Sanger, não têm o grupo hidroxila da ribose presente no DNA.

meiro, desnaturamos os dois filamentos desse segmento. Em seguida, criamos um primer para a síntese de DNA que irá hibridizar exatamente com um local no segmento de DNA clonado, e então adicionamos um “coquetel” de DNA polimerase, trifosfatos de desoxinucleotídeos normais (dATP, dCTP, dGTP e dTTP) e uma pequena quantidade de um dideoxinucleotídeo especial para uma das quatro bases (p. ex., trifosfato de dideoxiadenosina, abreviada ddATP). A polimerase começará a sintetizar o filamento complementar de DNA, começando a partir do primer, mas irá parar em qualquer ponto no qual o trifosfato de dideoxinucleotídeo é incorporado à cadeia crescente de DNA em lugar do trifosfato de desoxinucleotídeo normal. Suponha que a seqüência de DNA do segmento que estamos tentando seqüenciar é

5' ACGGGATAGCTAATTGTTTACCGCCGGAGCCA 3'

Começaríamos então a síntese de DNA a partir de um primer complementar:

5' ACGGGATAGCTAATTGTTTACCGCCGGAGCCA 3'
3' CGGCCTCGGT 5'

← Sentido da síntese de DNA

Usando a síntese especial do “coquetel” de DNA na ponta com ddATP, por exemplo, criaremos um conjunto de fragmentos de DNA que têm o mesmo início mas diferentes pontos finais, porque os fragmentos param em qualquer ponto da inserção de ddATP em vez de dATP parar a replicação de DNA. O arranjo de diferentes cadeias de DNA paradas por ddATP é parecido com a lista de seqüências a seguir. (*A indica o dideoxinucleotídeo.)

ATGGGATAGCTAATTGTTTACCGCCGGAGCCA 3'	Clone molde de DNA
CGGCCTCGGT 5'	Primer para a síntese
← Sentido da síntese de DNA	
*ATGGCGGCCTCGGT 5'	Fragmento dideoxi 1
*AATGGCGGCCTCGGT 5'	Fragmento dideoxi 2
*AAATGGCGGCCTCGGT 5'	Fragmento dideoxi 3
*ACAAATGGCGGCCTCGGT 5'	Fragmento dideoxi 4
*AACAAATGGCGGCCTCGGT 5'	Fragmento dideoxi 5
*ATTAACAAATGGCGGCCTCGGT 5'	Fragmento dideoxi 6
*ATCGATTAAACAAATGGCGGCCTCGGT 5'	Fragmento dideoxi 7
*ACCCTATCGATTAAACAAATGGCGGCCTCGGT 5'	Fragmento dideoxi 8

Podemos gerar um arranjo de tais fragmentos para cada um dos quatro trifosfatos de dideoxinucleotídeos possíveis em quatro “coquetéis” separados (um com uma ponta ddATP, um com ddCTP, um com ddGTP e um com ddTTP). Cada um produzirá um arranjo diferente de fragmentos, sem dois coquetéis de ponta produzindo fragmentos do mesmo tamanho. Além disso, se adicionarmos os resultados de todos os quatro coquetéis, veremos que os fragmentos podem ser ordenados em tamanho, com os tamanhos aumentando em uma base de cada vez. As etapas finais do processo são:

1. Apresentar os fragmentos em ordem de tamanho usando eletroforese. Eles aparecerão em quatro colunas separadas chamadas de C, A, G e T.
2. Marcar os filamentos recém-sintetizados de modo que possam ser visualizados após terem sido separados de acordo com o tamanho por eletroforese em gel. Fazer isso marcando o primer, ou radioativa ou fluorescentemente (marcação de iniciação), ou o trifosfato de dideoxinucleotídeo individual (marcação de término).

Os produtos de tais reações de seqüenciamento dideoxi são mostrados na Figura 20.16. Esse resultado é uma

O método de seqüenciamento dideoxi

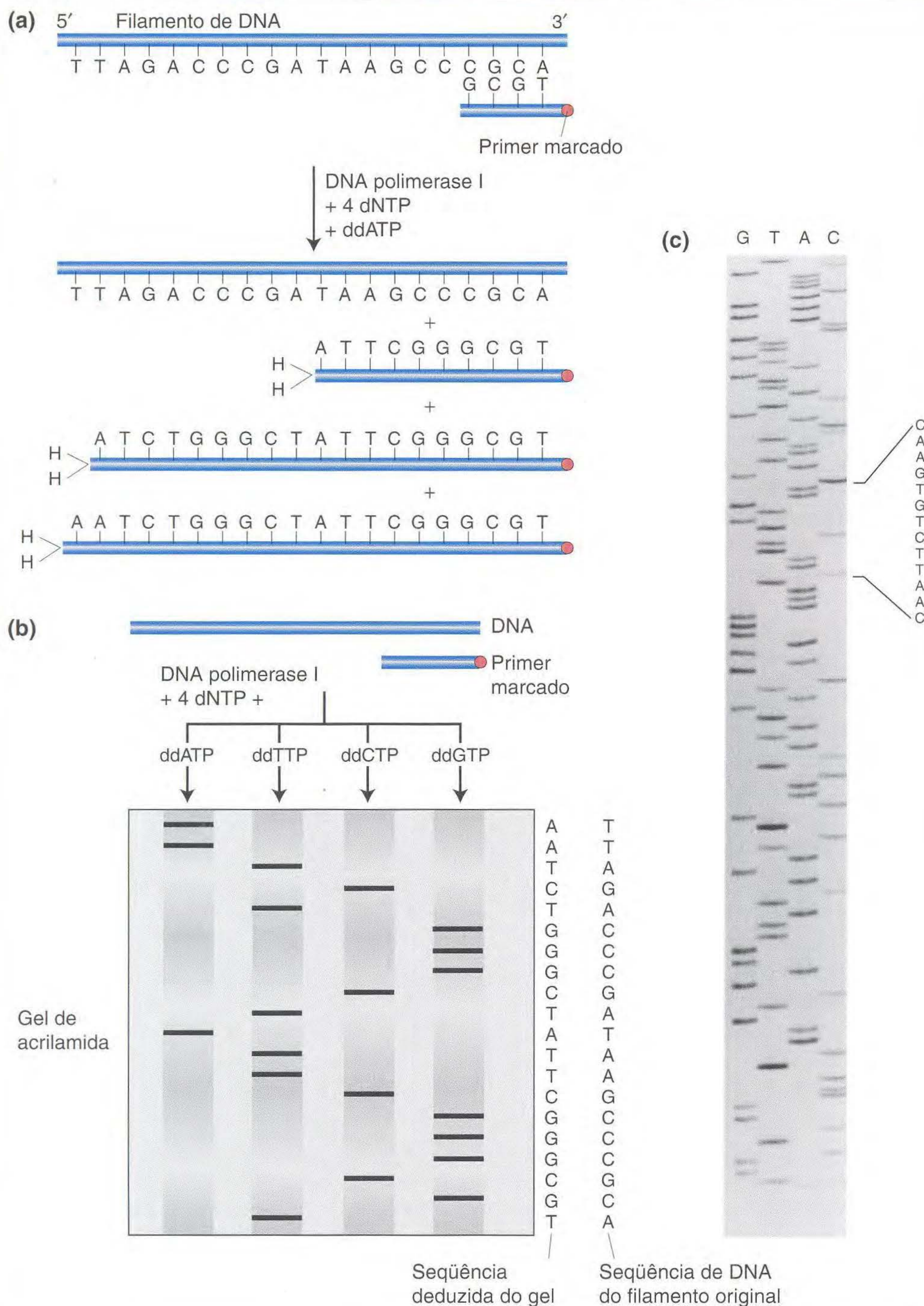


FIG. 20.16 O DNA é eficientemente seqüenciado pela inclusão de dideoxinucleotídeos entre os nucleotídeos usados para copiar um segmento de DNA. (a) Um primer (criado da seqüência flanqueadora do vetor) é usado para iniciar a síntese de DNA. A adição de quatro dideoxinucleotídeos diferentes (ddATP é mostrado aqui) pára aleatoriamente a síntese. (b) Os fragmentos resultantes são separados eletroforicamente e submetidos a auto-radiografia. A seqüência deduzida é mostrada à direita. (c) O gel de seqüenciamento de Sanger. [(a e b) de J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller, *Recombinant DNA*, 2nd ed. Copyright 1992 by Scientific American Books. (c) De Loida Escote-Carlson.]

escada de cadeias de DNA marcado aumentado de tamanho em um, e tudo o que precisamos é ler o gel para ler a sequência de DNA do filamento sintetizado no sentido de 5' para 3'.

Se a marcação é um corante fluorescente e é usado um emissor de cor fluorescente diferente para cada uma das quatro reações de ddNTP, então podem ocorrer quatro reações no mesmo tubo de ensaio, e os quatro conjuntos de cadeias de DNA podem passar por eletroforese juntos. Assim, podem ser produzidas quatro vezes mais seqüências ao mesmo tempo que podem ser produzidas fazendo as reações separadamente. Essa lógica é usada na detecção de fluorescência por máquinas de seqüenciamento automatizado de DNA. Graças a essas máquinas, o seqüenciamento de DNA pode ser produzido em grande escala, e as seqüências de todos os genomas podem ser obtidas com os procedimentos descritos nesta seção. A Figura 20.17 ilustra uma leitura de seqüenciamento automatizado. Cada pico colorido representa um fragmento de DNA de tamanho diferente, terminando com uma base fluorescente que foi detectada pelo *scanner* de fluorescência do seqüenciador automatizado de DNA; as quatro cores diferentes representam as quatro bases do DNA. As aplicações da tecnologia de seqüenciamento automatizado em uma escala genômica é o foco principal do Capítulo 13.

Um segmento clonado de DNA pode ser seqüenciado caracterizando um conjunto seriado de fragmentos truncados de DNA sintético, cada um terminando em posições diferentes correspondentes à incorporação de um didesoxinucleotídeo.

► 20.4 Análise Genética Direta Usando Clonagem Posicional

Considerando os inúmeros recursos genômicos que se acumulam rapidamente (p. ex., seqüências de genomas inteiros, mapeamento de populações, coleções de marcadores moleculares), torna-se cada vez mais possível isolar genes apenas com base no conhecimento de seu ponto em um mapa genético. A estratégia geral é chamada de *clonagem posicional*, e é ilustrada nesta seção pelo estudo de dois casos.

O poder da clonagem posicional é que ou mutantes ou variantes naturais podem ser usados como pontos de início na descoberta de genes. Em muitos casos, um geneticista deseja estudar um genoma em relação a todos os genes que contribuem para um processo biológico particular — p. ex., o desenvolvimento do cérebro. Em tais casos, após o geneticista mutagenizar uma grande população de genomas, o desafio é selecionar nessa coleção de indivíduos e identificar alguns com fenótipos sugestivos de uma mutação que afete o processo de interesse. A etapa seguinte é determinar uma posição aproximada de mapa do gene mutante como primeiro passo para isolar o gene. Após essa posição aproximada ter sido determinada, o geneticista está pronto para aplicar a clonagem posicional para determinar o local exato do gene e, finalmente, sua seqüência. Com essa informação disponível, o geneticista pode geralmente descobrir o que está errado com o gene mutante e qual é a função normal do gene tipo selvagem. A clonagem posicional também tem sido uma estratégia efetiva para compreender a base molecular da variação natural. No exemplo apresentado mais adiante nesta seção, você aprenderá como a clonagem posicional está sendo usada para identificar os genes selecionados pelos

Leitura da seqüência de DNA de um seqüenciador automático

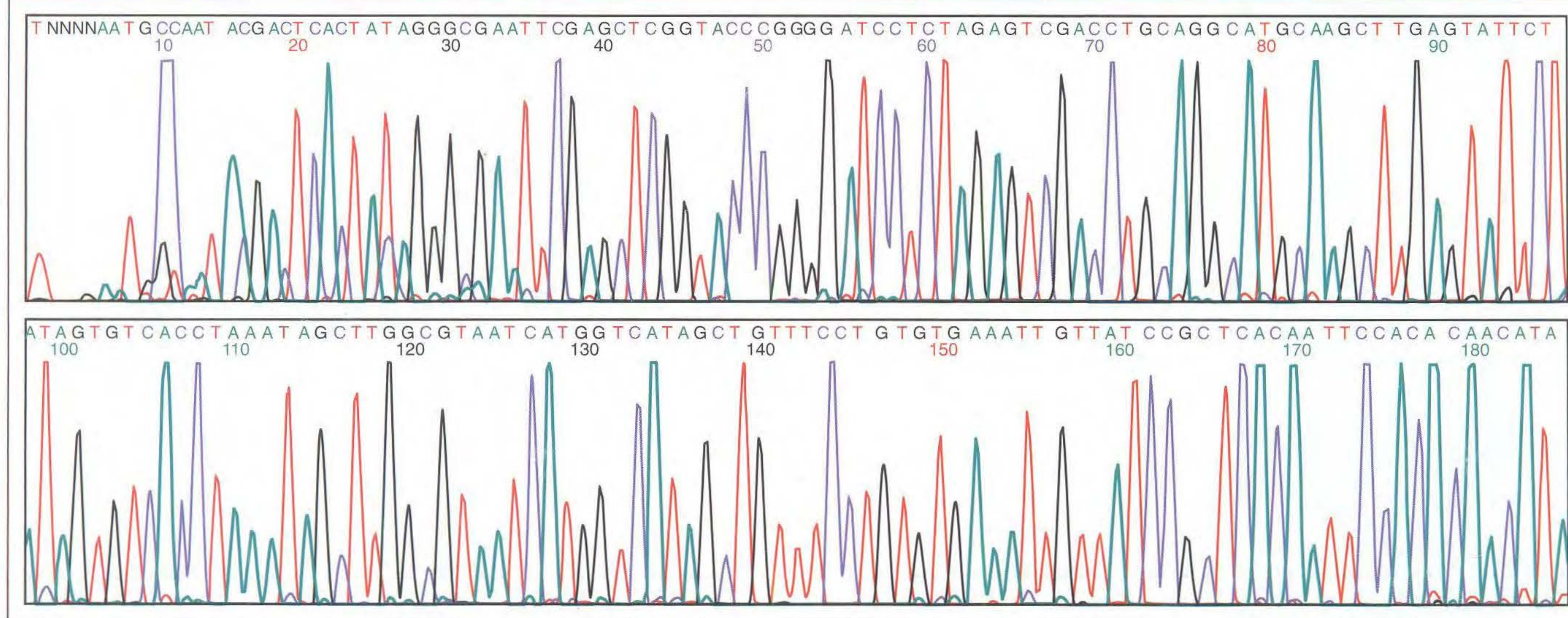


FIG. 20.17 Impressão de um seqüenciador automático que usa corantes fluorescentes. Cada uma das quatro cores representa uma base diferente. A letra "N" representa uma base que não pode ser atribuída, porque os picos estão muito baixos. Note que, se isso fosse um gel, como na Figura 20.16c, cada um desses picos corresponderia a uma das bandas escuras no gel. Em outras palavras, esses picos coloridos representam uma leitura diferente do mesmo tipo de dados que são produzidos no seqüenciamento de um gel.

americanos nativos na domesticação de milho a partir de seu genitor selvagem teosinto. Esses casos podem ser vistos como **genética direta**, porque o geneticista analisa fenótipos hereditáveis em âmbito genético antes de fazer a análise molecular de mutantes ou variantes.

Com as técnicas descritas neste capítulo e em outros, podemos ver como a análise direta é feita, da triagem de mutantes até a identificação molecular de um gene.

1. Tipicamente, a genética direta começa com o genoma tipo selvagem, que é mutado aleatoriamente com um mutágeno e sistematicamente avaliado quanto a mutações que têm algum fenótipo em comum. Esse procedimento às vezes é chamado de *caça ao mutante*. Idealmente, identifica mutações em literalmente todo o fenótipo. Então, o genoma pode ser dito como tendo sido *saturado* para mutações dessa classe. Por exemplo, a triagem de mutantes foi uma primeira etapa importante no isolamento dos genes ferramentais (*toolkit*) de *Drosophila* (veja Capítulo 12).
2. Os mutantes são identificados como tendo sido causados por mutações monogênicas, como descrito no Capítulo 2.
3. O gene é mapeado em uma posição aproximada no cromossomo, com o uso de técnicas descritas no Capítulo 4.
4. Marcadores moleculares ligados ao gene são usados para isolar clones de uma biblioteca existente ou de uma biblioteca que foi feita para esse propósito.
5. Como os genomas de organismos superiores são geralmente muito grandes (o genoma humano, por exemplo, tem quase 3.000 MB), os marcadores moleculares ligados disponíveis podem ainda ter mais de 1 milhão de pares de bases do gene de interesse. Nesse caso, os clones distantes servem como ponto de início para obter marcadores moleculares adicionais na região de interesse (veja Capítulo 4).
6. A sequência de DNA do clone selecionado é pesquisada quanto a genes candidatos (ORF).
7. Uma variedade de técnicas podem ser usadas para determinar qual ORF é o gene-alvo.

Essas etapas aplicam-se esteja ou não disponível uma sequência genômica anotada para o organismo em estudo. Entretanto, o processo é muito mais rápido e simples quando a sequência genômica é conhecida, como ilustra o estudo de caso na seção seguinte.

Uma análise direta para identificar um gene de doença humana

Vamos seguir os métodos para identificar a sequência genômica do gene de fibrose cística (FC). Nenhum defeito bioquímico primário era conhecido na época em que o gene foi isolado, e, assim, era mais um gene à procura de uma função.

As triagens genéticas podem ser usadas para dissecar qualquer processo biológico. Sua efetividade depende apenas da ingenuidade do pesquisador em criar um protocolo que revele a classe desejada das mutações. Entretanto, as triagens genéticas não podem ser usadas com seres humanos, porque não queremos intencionalmente criar mutantes huma-

nos. Assim, as análises de heredogramas de grandes famílias com a característica da doença são feitas (a análise de heredogramas é descrita no Capítulo 2) quando tal informação está disponível para determinar a posição do defeito genético causador da doença, tal como a fibrose cística ou doença de Huntington. Membros de uma família que tenham a doença são encontrados tendo um ou mais marcadores moleculares em comum que não são encontrados em outras famílias (veja marcadores moleculares no Capítulo 4). A localização dos marcadores fornece o local aproximado do gene (etapa 3 na lista precedente).

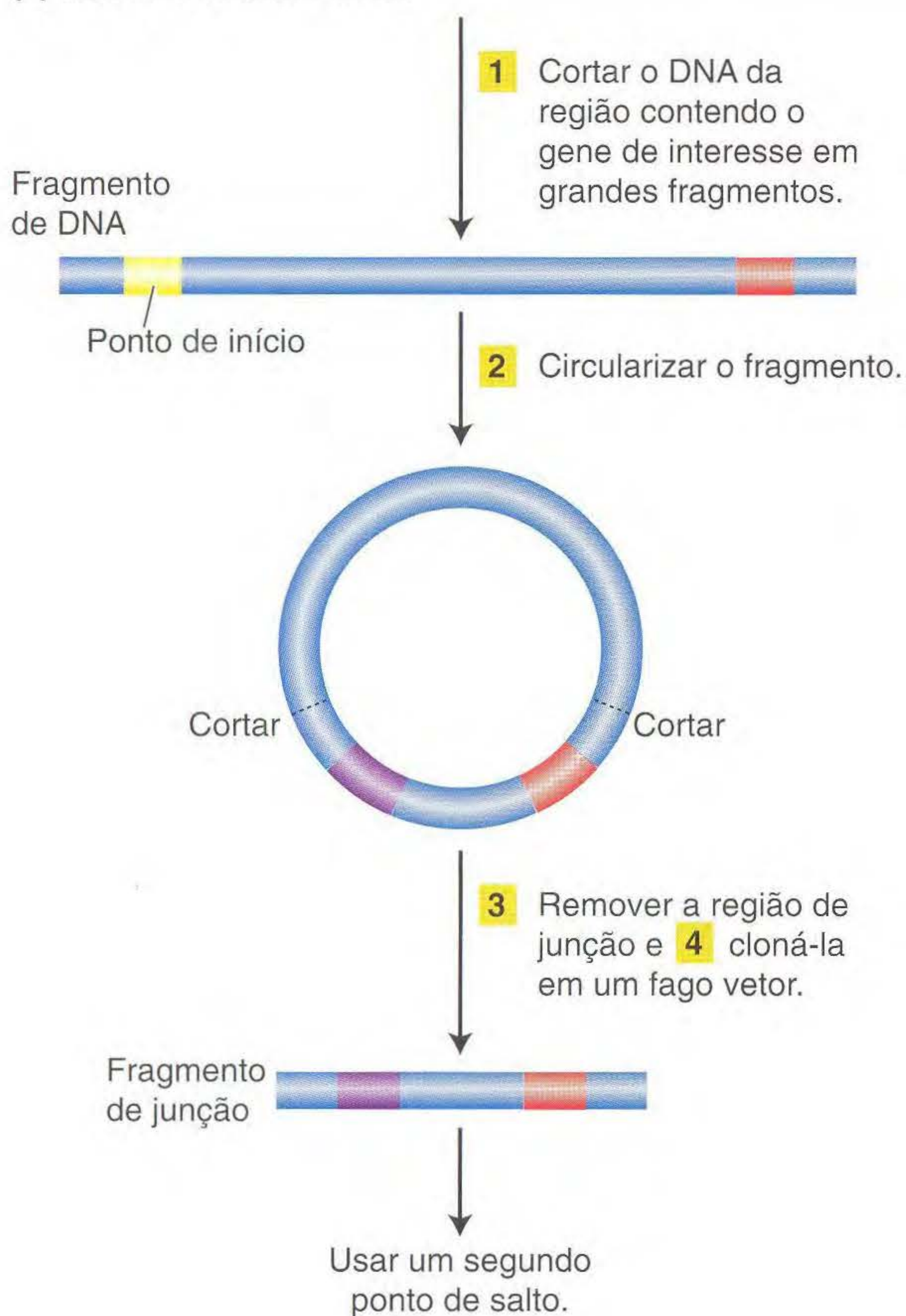
A ligação a marcadores moleculares situou o gene de FC no braço longo do cromossomo 7, entre as bandas 7q22 e 7q31.1. Pensou-se que o gene FC estivesse dentro dessa região, flanqueado pelo gene *met* (um proto-oncogene; veja Capítulo 15) em uma ponta e um marcador molecular, D788, na outra ponta. Mas, entre esses marcadores, há 1,5 centimorgans (unidades de mapa) de cromossomo, uma grande área não-mapeada com mais de 1 milhão de bases. Para chegar mais próximo, foi necessário gerar mais marcadores moleculares dentro dessa região. O método geral para isolamento de marcadores moleculares (descritos no Capítulo 4) é identificar uma região do DNA que é polimórfica em indivíduos ou populações que diferem na característica de interesse. Encontrando marcadores moleculares ligados ao gene FC, os geneticistas estreitaram a região contendo o gene FC para cerca de 500 kpb, ainda uma distância considerável.

Foi criado um mapa físico de toda a região; isto é, um conjunto aleatório de clones dessa região foi colocado na ordem correta. O ordenamento desses clones requer duas técnicas capazes de atravessar a enorme distância genética: o *andar no cromossomo* e uma técnica correlata chamada *saltar no cromossomo* (*chromosome jumping*). Esta última técnica fornece um modo de saltar sobre áreas potencialmente não-clonáveis de DNA e gerar marcos bem espaçados ao longo da sequência que podem ser usados como pontos iniciais para várias caminhadas bidirecionais no cromossomo. Essas caminhadas bidirecionais no cromossomo são usadas para ordenar os clones, como mostrado na Figura 20.13. O saltar no cromossomo é mostrado na Figura 20.18. O clone contendo moléculas marcadoras que são mais fortemente ligadas à característica FC é então sequenciado, e a caçada por alguns genes ao longo desse trecho de genes e sequências não-codificantes pode começar.

Genes candidatos foram identificados por características notáveis, tais como ilhas de CpG e sinais de fim, comuns a genes (veja Capítulo 13). No exemplo da FC, as sequências de genes candidatos e cDNA foram comparadas entre pessoas normais e pacientes com FC. Uma comparação da sequência de um cDNA com esse gene candidato em pessoas normais e pacientes com FC revelou uma deleção de três pares de bases nos pacientes com FC, eliminando uma fenilalanina da proteína. Por sua vez, a partir dessa sequência deduzida, a estrutura tridimensional da proteína foi prevista. Essa proteína é estruturalmente similar a proteínas de transporte de íons em outros sistemas, sugerindo que um defeito de transporte é a causa primária da FC. Quando usado para

Andar e salto no cromossomo

(a) Salto no cromossomo



(b) Andar no cromossomo a partir de pontos de salto

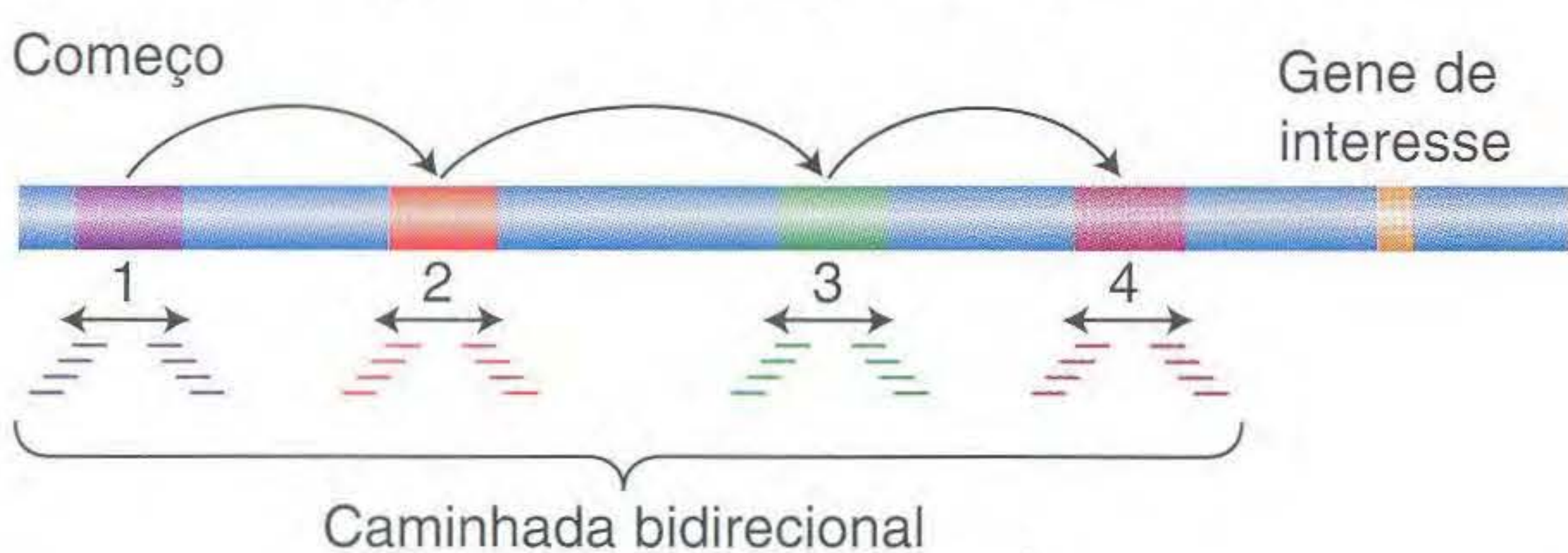


FIG. 20.18 Manipulando fragmentos genômicos clonados para saltar no cromossomo, um tipo modificado de andar no cromossomo que pode ultrapassar regiões difíceis de clonar. (a) Após todos os fragmentos terem sido clonados, como mostrado na parte a, uma sonda do começo da região do DNA em investigação (amarelo) é usada para triar os clones para encontrar qual o que contém o começo da sequência. Quando esse clone é encontrado, a outra ponta da sequência de junção (púrpura) é removida e usada para triar a biblioteca novamente para fazer um segundo salto. (b) De cada posição de salto, pode ser feita uma caminhada no cromossomo, em ambas as direções, para ordenar um segundo grupo de fragmentos clonados menores.

transformar linhagens celulares mutantes de pacientes com FC, o gene tipo selvagem restaurou o funcionamento normal; essa “recuperação” fenotípica foi a confirmação final de que a sequência isolada era de fato o gene FC.

Os protocolos de andar e saltar no cromossomo ainda são estratégias úteis para a clonagem posicional em organismos para os quais a sequência do genoma não está disponível. Esse era o caso do genoma humano em 1989, quando o gene para FC foi isolado. Entretanto, a disponibilidade da sequência genômica completa encurtou muito o tempo necessário para identificar um gene humano de doença, tornando obsoletas as técnicas que consomem muito tempo, tais como andar e saltar no cromossomo. Em vez disso, as buscas ajudadas por computador dos bancos de dados de informações do genoma humano servem para identificar clones portadores de marcadores moleculares e regiões adjacentes. Além disso, esses bancos de dados incluem informações sobre o conteúdo de genes dessas regiões (veja Capítulo 13, quanto a anotações), permitindo aos geneticistas identificar um ou mais possíveis genes candidatos. Um desses genes pode ser identificado como o gene de interesse usando os mesmos tipos de técnicas usadas para identificar o gene FC — por exemplo, comparando sequências de genes candidatos e pacientes com FC.

Uma análise direta para identificar um gene importante para a domesticação do milho

O QUE OS GENETICISTAS ESTÃO FAZENDO HOJE

Como mencionado no começo do capítulo, George Beadle propôs que, a despeito de marcantes diferenças morfológicas, o teosinto selvagem é o ancestral do milho moderno domesticado. Fazendo milhares de cruzamentos genéticos entre o milho e o teosinto (eles são interférteis), Beadle concluiu que apenas cinco loci de características quantitativas (QTL, veja Capítulos 3 e 18) de efeito importante podem responder por suas diferenças morfológicas (Figura 20.19).

O isolamento de genes que transformam teosinto em milho foi de grande interesse, especialmente para o geneticista John Doebley. Ele escreveu que a etapa mais crítica na domesticação do milho era “a liberação do grão do envoltório protetor que reveste o grão no teosinto”. O envoltório duro permite que as sementes de teosinto sejam dispersas após passar ilesas pelo trato digestivo de um animal faminto. Entretanto, para ser útil como fonte de alimento, a planta precisa ter uma semente exposta sem o envoltório. O QTL responsável pelos “grãos nus de milho” é chamado de *gluma de arquitetura do teosinto* (*tga1*). Com base em cruzamentos genéticos, Doebley e seus colaboradores supuseram que o alelo de milho de *tga1* (chamado de *Tga1*-milho) promove grãos nus, enquanto o alelo de teosinto (chamado de *tga1*) promove grãos duros.

Para isolar o gene *tga1*, Doebley teve primeiro que gerar uma linhagem de milho que continha a região do cromossomo de teosinto com o alelo de teosinto (chamado milho:*tga1*). Nessa situação, o alelo de teosinto é dito introgridido no ambiente do milho. [Similarmente, a região do cromossomo do milho contendo *Tga1* pode ser introgridida no ambiente do teosinto (teosinto:*Tga1*-milho); Figura 20.20.] A linhagem de milho:*tga1* então serviu como o ponto de início de uma estratégia de clonagem posicional. Em termos gerais, a linhagem milho:*tga1* difere do milho apenas por conter um

Principais diferenças morfológicas entre teosinto selvagem e milho domesticado



FIG. 20.19 Um teosinto selvagem tem relativamente poucos grãos, todos em envoltórios duros, enquanto no milho domesticado há muito mais grãos, todos expostos. [Cortesia de Nicolle Roger Fuller, National Science Foundation.]

pedaço do cromossomo de teosinto com seu alelo *tga1* (e vários outros genes). Comparando os marcadores genéticos nas duas linhagens, Doebley foi capaz de isolar as diferenças entre eles, isto é, o fragmento do cromossomo de teosinto contendo vários genes de teosinto, incluindo *tga1*. Como o último estágio da maioria das estratégias de clonagem posicional, essa consiste em determinar qual dos vários genes na região era *tga1*. E, como a maioria das estratégias de clonagem posicional, essa determinação era baseada em uma análise

comparativa das seqüências dos alelos de milho e teosinto. Essa análise acabou levando Doebley e seus colaboradores a identificar *tga1* como um suposto fator regulador de transcrição (veja Capítulo 11) que diferia em milho e teosinto por algumas mudanças de pares de bases, incluindo uma única diferença de aminoácidos. Além disso, eles determinaram que as linhagens com *tga1* tinham mais proteína TGA1 na espiga em desenvolvimento do que as linhagens *tga1*-milho. Eles prosseguiram para demonstrar que essa diferença era devida a uma diferença em regiões reguladoras do gene que estavam situadas há mais de 40 kb (40.000 pares de bases) antecedentes ao ponto de início da transcrição. Esse achado os levou a supor que diferenças complexas entre as sementes envolturadas de teosinto e os grãos nus de milho resultam da expressão de variantes alélicas diferentes de *tga1*. Isto é, o gene de milho é provavelmente não um alelo nulo de *tga1*, mas uma forma alternativa do gene que leva a uma forma alternativa do desenvolvimento da espiga.

As seções precedentes introduziram as técnicas fundamentais que revolucionaram a Genética. As duas últimas seções deste capítulo enfocarão a aplicação dessas técnicas ao diagnóstico de doenças humanas e consulta genética.

Resultados de inserção de um gene de milho em teosinto



FIG. 20.20 Uma espiga de teosinto com grãos típicos (esquerda) é mostrada perto de uma espiga de teosinto com o alelo *tga1* de milho de introgressão (milho:*tga1*). Note os grãos expostos (direita). [Cortesia de J.F. Doebley, Department of Genetics, University of Wisconsin—Madison.]

► 20.5 Detecção de Alelos de Doenças Humanas: Diagnóstico Genético Molecular

Um fator contribuinte em mais de 500 doenças genéticas humanas, incluindo alcaptonúria, é um alelo mutante recessivo de um único gene. Para famílias em risco de tais doenças, é importante detectar genitores prospectivos heterozigotos

para permitir a consulta* apropriada. Também é necessário ser capaz de detectar bem cedo uma prole homozigota, idealmente no estágio fetal, de modo que os médicos possam aplicar logo terapias com drogas ou dietas. No futuro, pode até haver a possibilidade de terapia gênica. Distúrbios dominantes também podem necessitar de diagnóstico genético. Por exemplo, as pessoas em risco de doença de Huntington de manifestação tardia precisam saber se possuem o alelo da doença antes de terem filhos.

Testes amplamente utilizados são capazes de detectar alelos defeituosos homozigotos em células fetais. As células fetais podem ser colhidas do líquido amniótico, separadas de outros componentes, e cultivadas para permitir a análise de cromossomos, proteínas, reações enzimáticas e outras propriedades bioquímicas. Esse processo, a **amniocentese** (Figura 20.21), pode identificar vários distúrbios conhecidos. O Quadro 20.1 mostra alguns exemplos. A **punção de vilosidades coriônicas (CVS)** é uma técnica correlata na qual uma pequena amostra de células da placenta é aspirada com uma longa seringa. A CVS pode ser feita mais cedo na gestação do que a amniocentese, que deve esperar o desenvolvimento de um volume grande o suficiente de líquido amniótico.

Tradicionalmente, esses procedimentos de triagem só têm identificado distúrbios que podem ser detectados como um defeito químico nas células cultivadas. Entretanto, com a tecnologia do DNA recombinante, o DNA pode ser analisado diretamente. Em princípio, o gene fetal apropriado pode ser clonado e sua sequência comparada com a de um gene normal clonado para ver se o gene fetal é normal. Entretanto,

*N.T.: *Counseling* foi traduzida como consulta genética, e não aconselhamento, pois não se deve dar conselhos, mas sim informações que permitam ao casal tomar uma decisão bem informada.

Quadro 20.1 Algumas Doenças Genéticas Comuns

Erros inatos do metabolismo	Incidência aproximada entre nativos
1. Fibrose cística (defeito da proteína do canal de cloreto)	1/1.600 caucasianos
2. Distrofia muscular de Duchenne (proteína muscular defeituosa)	1/3.000 meninos (ligada ao X); distrofina
3. Doença de Gaucher (defeito de glicocerebrosidase)	1/2.500 judeus Asquenazi; 1/75.000 outros
4. Doença de Tay-Sachs (defeito de hexosaminidase A)	1/3.500 judeus Asquenazi; 1/35.000 outros
5. Pentosúria essencial (uma condição benigna)	1/2.000 judeus Asquenazi; 1/50.000 outros
6. Hemofilia clássica (defeito de fator VIII de coagulação)	1/10.000 meninos (ligada ao X)
7. Fenilcetonúria (defeito de fenilalanina hidroxilase)	1/5.000 irlandeses celtas; 1/15.000 outros
8. Cistinúria (defeito de transporte de membrana de cistina)	1/15.000
9. Leucodistrofia metacromática (defeito de arilsulfatase A)	1/40.000
10. Galactosemia (defeito de galactose-1-fosfato uridil transferase)	1/40.000
11. Anemia falciforme (defeito de cadeia β de globina)	1/400 negros dos EUA
Em algumas populações africanas, a frequência de heterozigotos é de 40%	
12. Talassemia (redução ou ausência de cadeia de globina)	1/400 entre algumas populações do Mediterrâneo

Nota: Embora a grande maioria das mais de 500 doenças genéticas recessivas reconhecidas sejam extremamente raras, em combinação elas constituem uma enorme carga de sofrimento humano. Como é consistente nas mutações mendelianas, a incidência de algumas dessas doenças é muito mais alta em alguns grupos raciais do que em outros.

Fonte: J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller, *Recombinant DNA*, 2nd ed. Copyright 1992 by Scientific American Books.

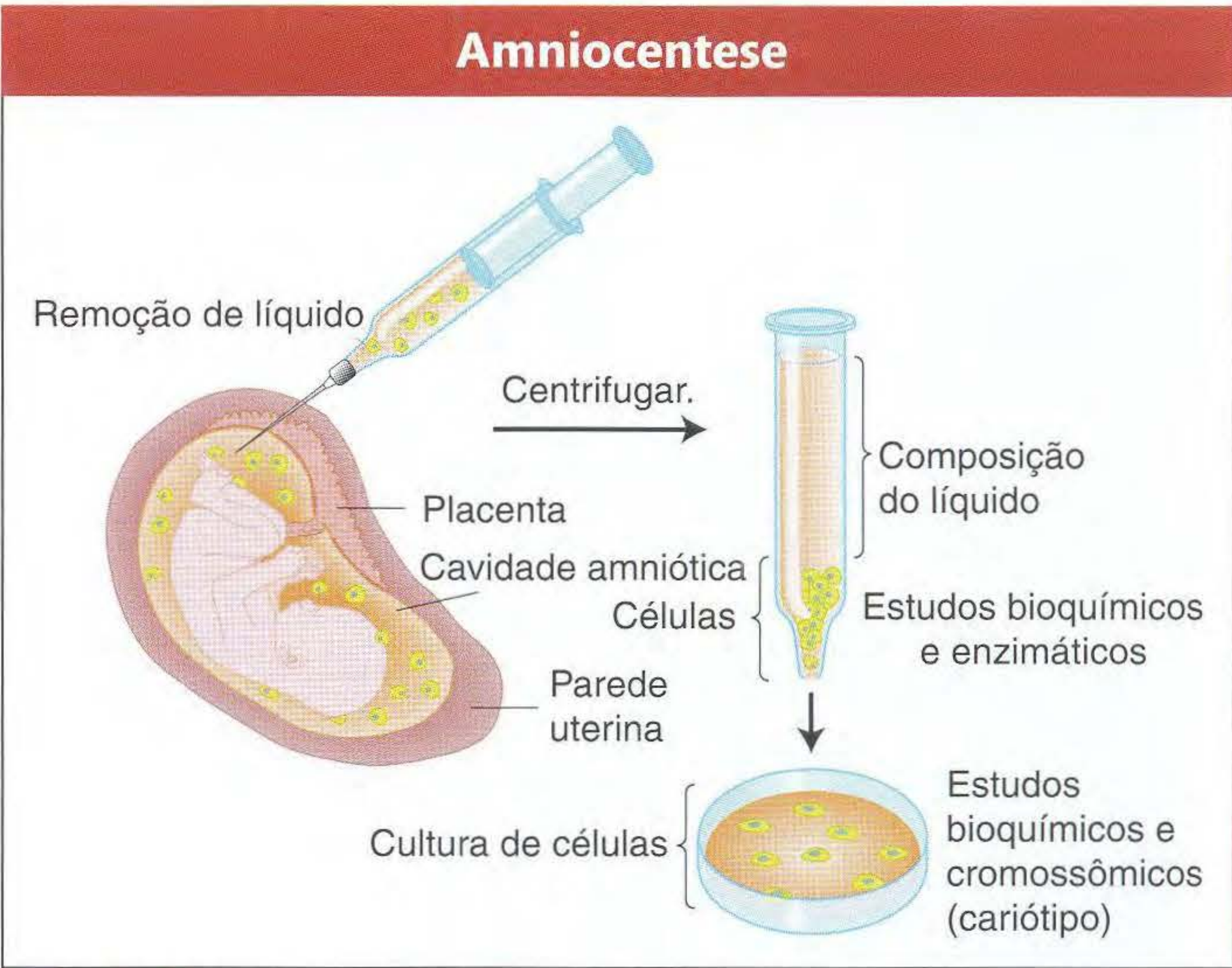


FIG. 20.21 Na amniocentese, as células colhidas do líquido amniótico são analisadas para detectar alelos defeituosos homozigotos.

esse procedimento seria demorado e impraticável, razão pela qual foram criados atalhos. Por exemplo, como a PCR permite que o pesquisador situe qualquer sequência desejada, ela pode ser usada para amplificar e, depois, seqüenciar qualquer sequência de DNA potencialmente defeituosa. Em um enfoque ainda mais simples, podem ser criados primers para se hibridizar com o alelo normal e, portanto, iniciar sua amplificação, mas não se hibridizar ao alelo mutante. Essa

técnica pode diagnosticar doenças causadas pela presença de um sítio mutacional específico.

A tecnologia de DNA recombinante fornece muitas técnicas sensíveis para testar alelos defeituosos.

► 20.6 Engenharia Genética

Graças à tecnologia do DNA recombinante, os genes podem ser isolados em um tubo de ensaio e caracterizados como seqüências específicas de nucleotídeos. Mas, mesmo essa conquista não é o final da história. Veremos, em seguida, que o conhecimento de uma seqüência é geralmente o começo de uma nova rodada de manipulação. Quando caracterizada, uma seqüência pode ser manipulada para alterar o genótipo de um organismo. A introdução de um gene alterado em um organismo tornou-se um aspecto central da pesquisa de genética básica, mas também tem aplicação comercial. Dois exemplos desta última são (1) cabras podem secretar em seu leite antibióticos derivados de um fungo e (2) plantas não ficam congeladas pela incorporação de genes “anticongelantes” de um peixe do Ártico em seus genomas. O uso de técnicas de DNA recombinante para alterar o genótipo e o fenótipo de organismos desse modo é chamado de *engenharia genética*.

As técnicas de engenharia genética descritas na primeira parte deste capítulo foram originalmente desenvolvidas em bactérias e precisaram ser ampliadas para o modelo de eucariontes, que constituem uma grande proporção dos organismos-modelo de pesquisa. Os genes eucarióticos ainda são tipicamente clonados e seqüenciados em hospedeiros bacte-

rianos, mas ocasionalmente são introduzidos em um eucarionte, seja a espécie original doadora ou uma completamente diferente. O gene transferido é chamado de **transgene**, e o produto construído é chamado de **organismo transgênico**.

O transgene pode ser introduzido em uma célula eucariótica por uma variedade de técnicas, incluindo transformação, injeção, infecção bacteriana ou viral e bombardeamento com DNA revestido por tungstênio ou partículas de ouro (Figura 20.22). Quando o transgene entra em uma célula, ele é capaz de viajar para o núcleo, onde, para se tornar parte estável do genoma, deve inserir-se em um cromossomo ou (apenas em algumas espécies) replicar-se como parte de um plasmídeo. Caso ocorra a inserção, o transgene pode ou substituir o gene residente, ou inserir-se ectopicamente, isto é, em outros locais no genoma. Os transgenes de outras espécies tipicamente inserem-se ectopicamente.

A transgênese pode introduzir material genético novo ou modificado em células eucarióticas.

Veremos agora alguns exemplos em fungos, plantas, animais e tentativas de terapia gênica em humanos.

Engenharia genética em *Saccharomyces cerevisiae*

É correto dizer que *S. cerevisiae* é o modelo genético eucariótico mais sofisticado. A maioria das técnicas usadas para engenharia genética em geral foram desenvolvidas em levedura; assim, vamos considerar os caminhos gerais para transgênese em leveduras.

Os vetores mais simples de leveduras são plasmídeos integrativos (YIp), derivados de plasmídeos bacterianos nos quais o DNA da levedura de interesse foi inserido. Quando transformados em células de leveduras, esses plasmídeos inserem-se em cromossomos de levedura, geralmente por recombinação homóloga com o gene residente, por um crossing único ou duplo (Figura 20.23). Como resultado, ou todo o plasmídeo é inserido, ou o alelo-alvo é substituído pelo alelo do plasmídeo. O último é um exemplo de *substituição gênica* — nesse caso, a substituição é de um gene originalmente na célula de levedura por um gene construído. A substituição gênica pode ser usada para deletar um gene ou substituir um alelo mutante por sua contraparte tipo selvagem, ou, contrariamente, substituir um alelo tipo selvagem por um mutante. Tais substituições podem ser detectadas colocando as células em um meio que seleciona um alelo marcador no plasmídeo.

A origem bacteriana da replicação é diferente das origens eucarióticas, e, assim, os plasmídeos bacterianos não se replicam em leveduras. Portanto, o único modo pelo qual tais vetores podem gerar um genótipo estável modificado é serem integrados ao cromossomo de levedura.

Engenharia genética em plantas

Devido a seu significado econômico em agricultura, muitas plantas foram objeto de análises genéticas destinadas a desenvolver variedades aperfeiçoadas. A tecnologia do DNA

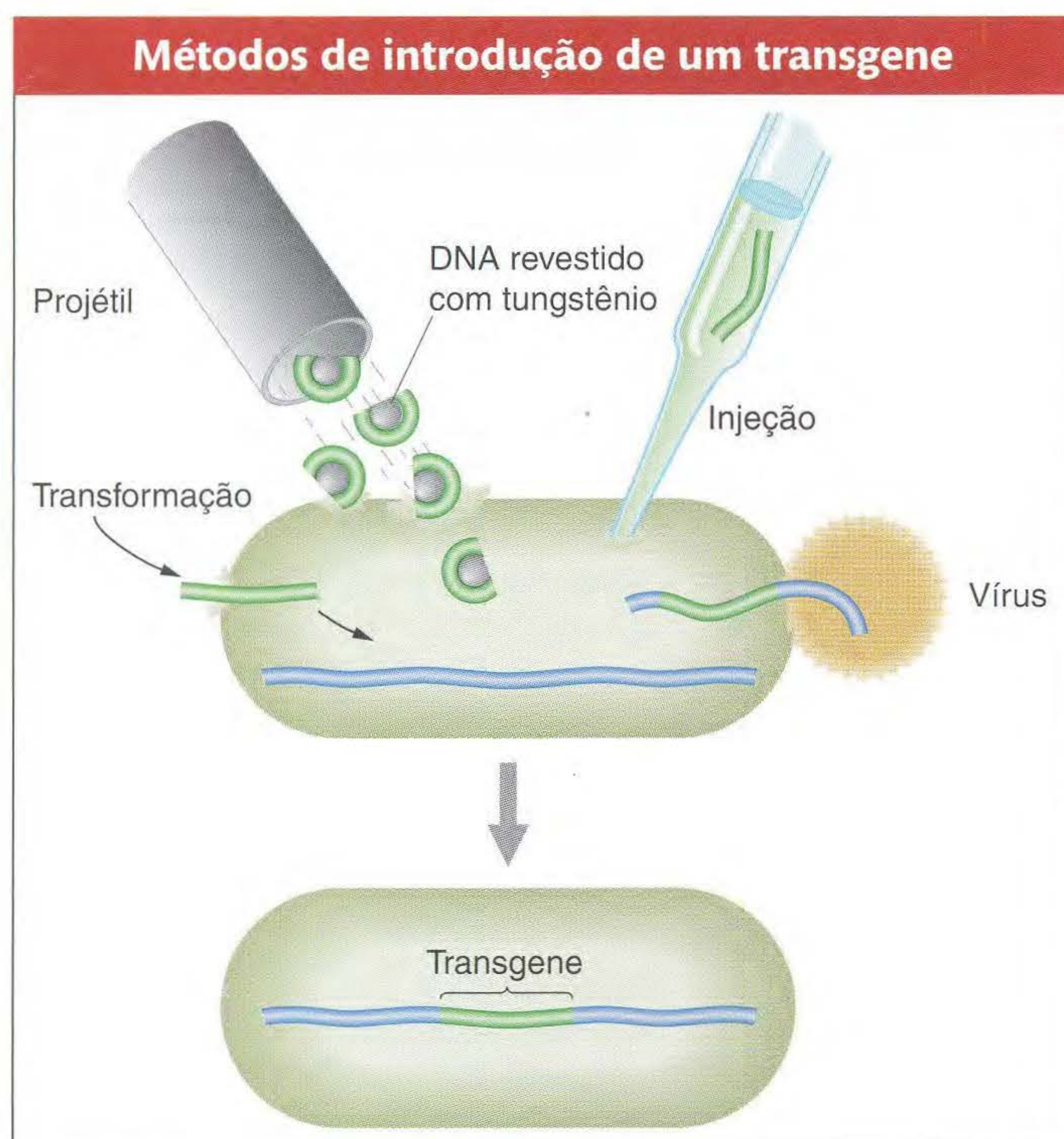


FIG. 20.22 Alguns dos modos diferentes de introduzir DNA exógeno em uma célula.

Dois resultados de transformação por vetores de leveduras simples

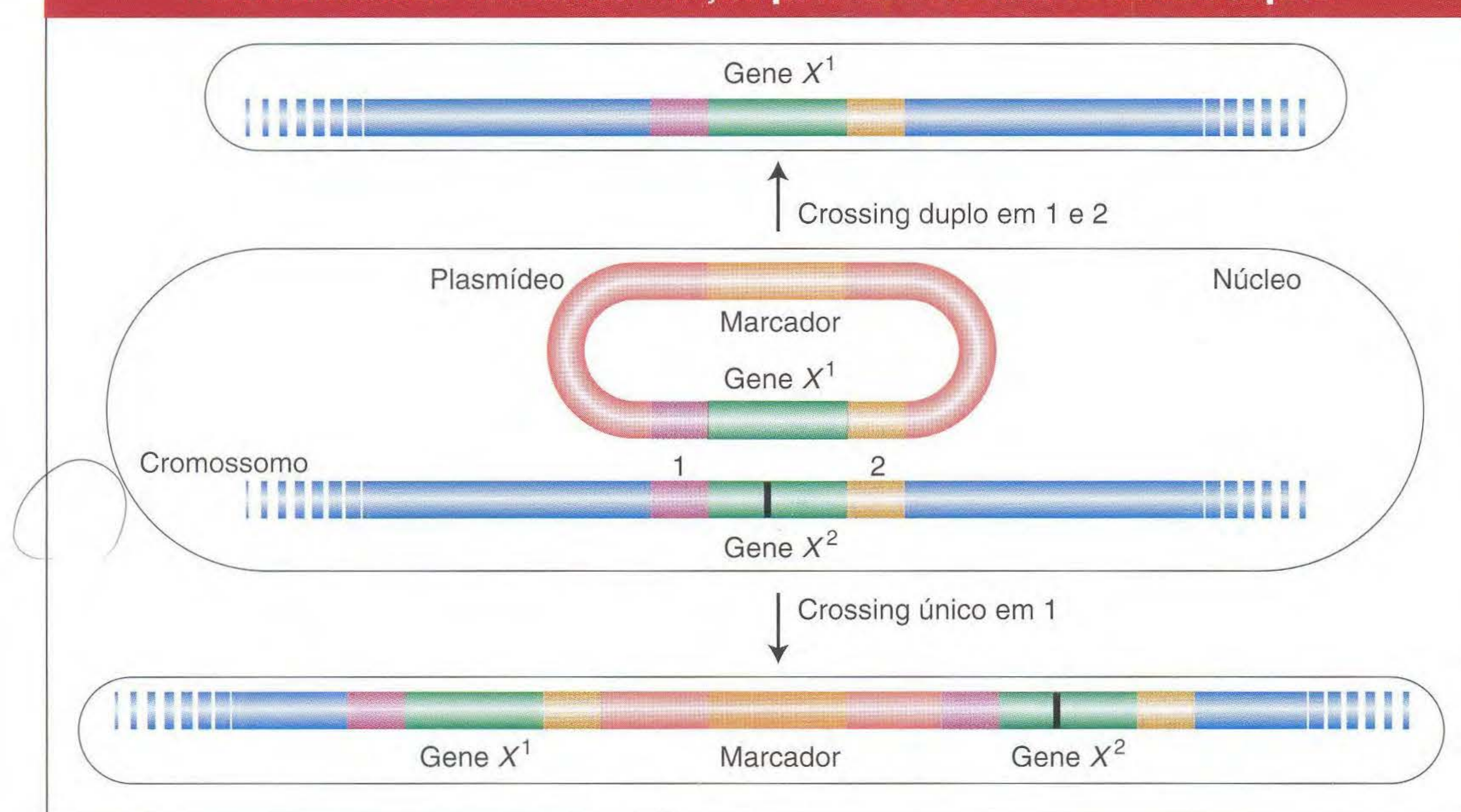


FIG. 20.23 Um plasmídeo possuindo um alelo ativo (gene X^+) insere-se em uma linhagem receptora de levedura portando um gene defeituoso (X^-) por recombinação homóloga. O resultado pode ser a substituição do gene defeituoso X^- (em cima) ou sua retenção juntamente com o novo alelo. O sítio mutante do gene X^- é representado como uma barra preta vertical. Crossings únicos na posição 2 também são possíveis, mas não são mostrados.

recombinante introduziu uma nova dimensão a esse esforço porque as modificações do genoma possibilitadas por essa tecnologia são quase ilimitadas. A diversidade genética não é mais obtida apenas selecionando variantes dentro de uma determinada espécie. O DNA agora pode ser introduzido de outras espécies de plantas, animais ou mesmo bactérias. Em resposta a novas possibilidades, um setor do público manifestou preocupação de que a introdução de **organismos geneticamente modificados (GMO)** nos alimentos talvez possa produzir problemas de saúde inesperados. As preocupações sobre os GMO representam uma faceta de um contínuo debate sobre saúde pública complexa, segurança, ética e aspectos educacionais criados pelas novas tecnologias genéticas.

Sistema de Plasmídeo Ti Um vetor rotineiramente usado para produzir plantas transgênicas é o **plasmídeo Ti**, um plasmídeo natural derivado da bactéria do solo chamada *Agrobacterium tumefaciens*. Essa bactéria causa o que é conhecido como *doença de galha (crown gall)*, na qual a planta infectada produz crescimentos descontrolados chamados tumores ou galhas. Estas galhas normalmente se formam na base (coroa) do tronco da planta. A chave para a produção do tumor é um grande plasmídeo de DNA circular (200 kb), o *plasmídeo Ti (indutor de tumor)*. Quando a bactéria infecta uma célula vegetal, uma parte do plasmídeo Ti é transferida e inserida, ao que parece mais ou menos aleatoriamente, no genoma da planta hospedeira (Figura 20.24). A região do plasmídeo Ti que se insere na planta hospedeira é chamada de *T-DNA* para DNA de transferência. A estrutura de um plasmídeo Ti é mostrada na Figura 20.25. Os genes cujos produtos catalisam essa transferência de T-DNA residem em uma região do plasmídeo Ti separada da própria região do T-DNA. A região do T-DNA codifica várias funções interessantes que contribuem para a habilidade do crescimento e divisão dentro da célula

da planta. Essas funções incluem enzimas que contribuem para a produção do tumor e outras proteínas que dirigem a síntese de compostos chamados *opinas*, substratos importantes para o crescimento da bactéria. Uma opina importante é a nopalina. As opinas de fato são sintetizadas pelas células vegetais infectadas, que expressam os genes sintetizadores de opinas situados na região T-DNA transferida. As opinas são importadas para as células bacterianas do tumor crescente e metabolizadas por enzimas codificadas pelos genes da bactéria que usam opina no plasmídeo Ti.

O comportamento natural do plasmídeo Ti torna-o bem adequado ao papel de um vetor para engenharia genética de plantas. Se o DNA de interesse puder ser inserido no T-DNA, então todo o conjunto seria inserido em um estado estável no cromossomo da planta. Esse sistema foi feito para funcionar essencialmente desse modo, mas com algumas modificações necessárias. Examinemos um protocolo.

Os plasmídeos Ti são muito grandes para serem facilmente manipulados e não podem ser prontamente reduzidos, pois contêm poucos sítios únicos de restrição, e porque grande parte do plasmídeo é necessária para sua replicação ou para a infecção e processo de transferência. Portanto, um plasmídeo Ti apropriadamente construído é criado em etapas. As primeiras etapas de clonagem ocorrem em *E. coli*, com o uso de um vetor intermediário consideravelmente menor que Ti. O próprio vetor intermediário se insere com o transgene no T-DNA. Esse vetor intermediário pode então ser recombinado com um segundo, plasmídeo Ti “desarmado”, formando um *plasmídeo cointegrado* que pode ser introduzido em uma célula vegetal por infecção de *Agrobacterium* e transformação. Um elemento importante no plasmídeo cointegrado é um marcador selecionável que pode ser usado para detectar células transformadas. A resistência à canamicina é um desses marcadores.

Infeção pelo plasmídeo Ti

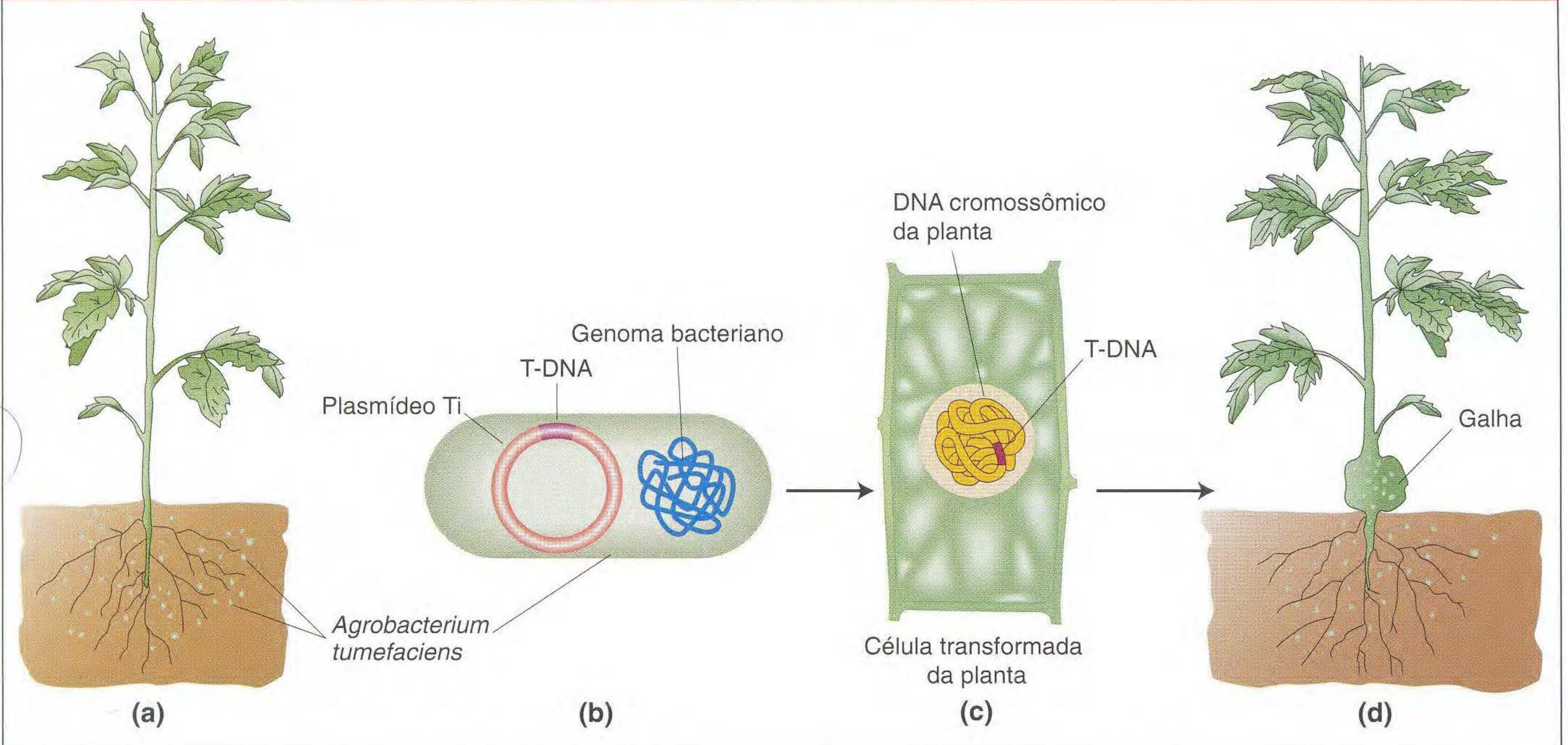


FIG. 20.24 No processo de causar a doença galha, a bactéria *Agrobacterium tumefaciens* insere uma parte de seu plasmídeo Ti, uma região chamada T-DNA, em um cromossomo da planta hospedeira.

Como mostra a Figura 20.26, as bactérias contendo um plasmídeo cointegrado são usadas para infectar segmentos cortados de tecido vegetal, tais como discos retirados de folhas. Em células infectadas, qualquer material genético entre seqüências flanqueadoras de T-DNA pode ser inserido em um cromossomo de planta. Se os discos de folhas são colo-

cados em um meio contendo canamicina, as únicas células vegetais que sofrerão divisão celular são as que adquiriram o gene *kan^R* introduzido no plasmídeo cointegrado. As células transformadas crescem em um aglomerado, ou calo, que pode ser induzido a formar brotos e raízes. Esses calos são transferidos para o solo, onde se desenvolvem em plantas transgênicas (veja Figura 20.26). Tipicamente, apenas uma única cópia da região de T-DNA insere-se em determinado genoma de planta, onde se segrega na meiose como um alelo mendeliano regular (Figura 20.27). A presença do inserto pode ser verificada por triagem do tecido transgênico quanto a marcadores genéticos transgênicos ou à presença de nopalina, ou pela triagem de DNA purificado com uma sonda de T-DNA em uma hibridização de Southern.

As plantas transgênicas levando qualquer uma de uma variedade de genes exógenos estão em uso atual, incluindo cultivos levando genes que conferem resistência a algumas pestes bacterianas ou por fungos, e muitas outras estão em desenvolvimento. Não só estão as qualidades das plantas em si sendo manipuladas, mas, como os microrganismos, as plantas também estão sendo usadas como “fábricas” convenientes para produzir proteínas codificadas por genes exógenos.

Engenharia genética em animais

As tecnologias transgênicas agora estão sendo empregadas com muitos sistemas-modelo animais. Enfocaremos três espécies animais mais intensamente usadas para pesquisa genética básica: o nematódeo *Caenorhabditis elegans*, a mosca-das-frutas *Drosophila melanogaster* e o camundongo *Mus musculus*. Versões de muitas das técnicas considera-

O plasmídeo Ti de *Agrobacterium tumefaciens*

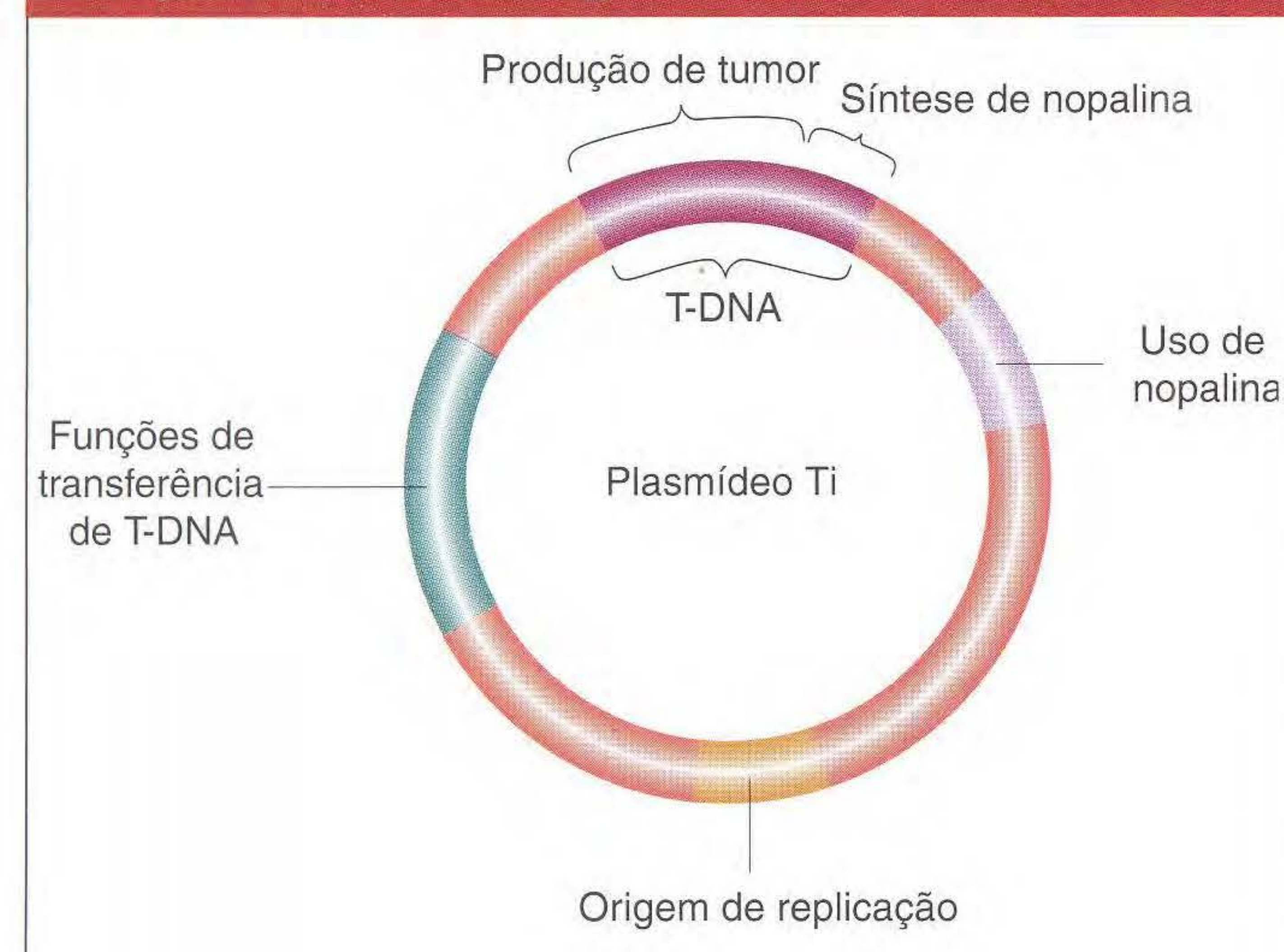


FIG. 20.25 Representação simplificada das principais regiões do plasmídeo Ti de *A. tumefaciens*. O T-DNA, quando inserido no DNA cromossômico da planta hospedeira, dirige a síntese de nopalina, que é então usada pela bactéria para seus próprios fins. T-DNA também dirige a célula vegetal para se dividir de modo descontrolado, produzindo um tumor.

A geração de uma planta transgênica

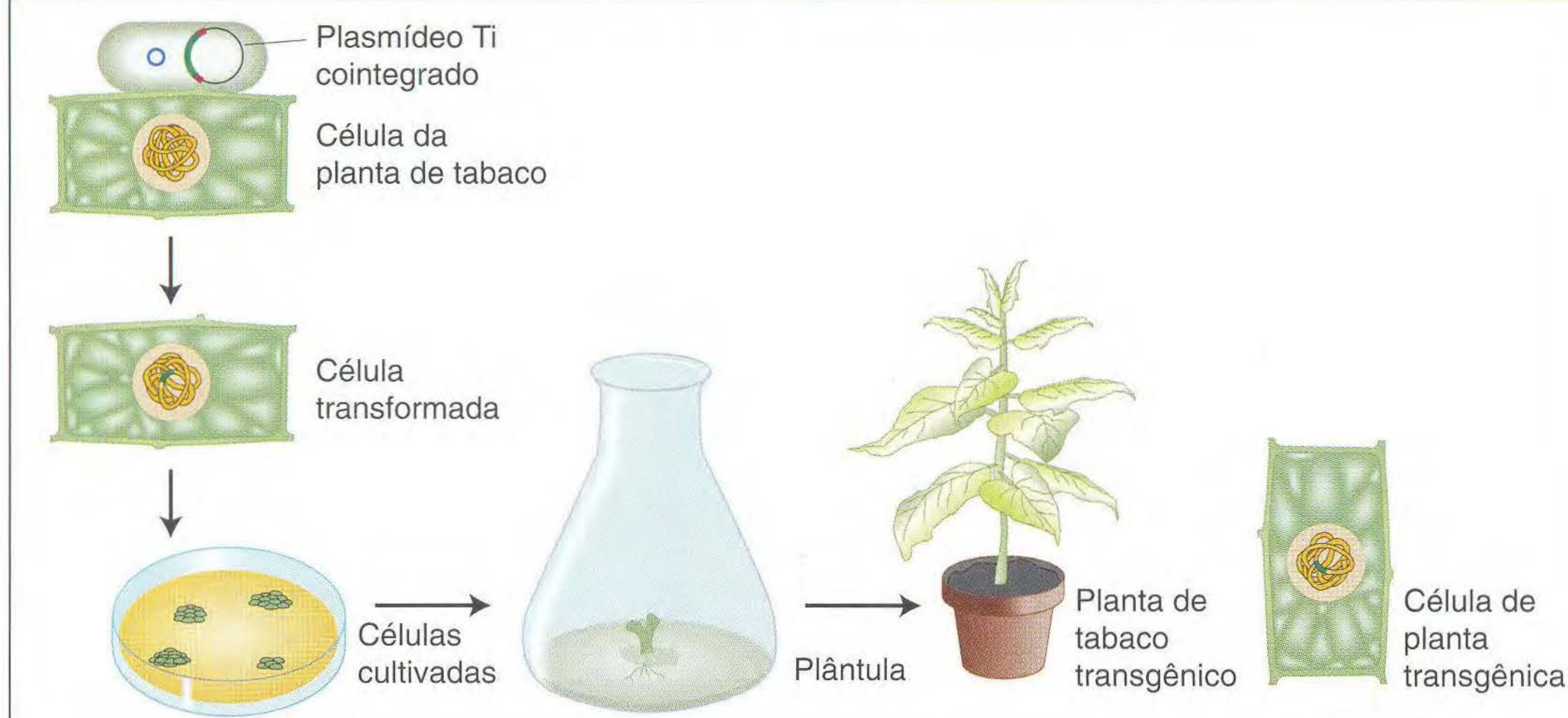


FIG. 20.26 Uma planta transgênica é gerada pelo crescimento de uma célula transformada pelo T-DNA.

das até agora também podem ser aplicadas nesses sistemas animais.

Transgênese em *C. elegans* O método usado para introduzir transgenes em *C. elegans* é simples: os DNA transgênicos são injetados diretamente no organismo, tipicamente como plasmídeos, cosmídeos ou outros DNA clonados em bactérias. A estratégia de injeção é determinada pela biologia reprodutiva do verme. As gônadas do verme são sinciciais, significando que existem muitos núcleos dentro da mesma célula gonadal. Uma célula sincicial é uma grande proporção de um braço da gônada, e a outra célula sincicial é o grosso do outro braço (Figura 20.28a). Esses núcleos não formam células individuais até a meiose, mas começam sua transformação em ovócitos individuais ou espermatozoides. Uma solução de DNA é injetada na região sincicial de um dos braços, expondo assim mais de 100 núcleos ao DNA transformante. Por acaso, alguns desses núcleos irão incorporar o DNA (lembre-se de que a membrana nuclear é rompida no curso da divisão, e, assim, o citoplasma no qual o DNA é injetado torna-se contínuo com o nucleoplasma). Tipicamente, o DNA transgênico forma arranjos extracromossômicos multicópia (Figura

20.28b) que existem como unidades independentes fora dos cromossomos. Mais raramente, os transgenes tornam-se integrados em uma posição ectópica em um cromossomo, ainda um arranjo multicópia. Infelizmente, as seqüências podem tornar-se embaralhadas dentro dos arranjos, complicando o trabalho do pesquisador.

Transgênese em *D. melanogaster* A transgênese em *D. melanogaster* requer uma técnica mais complexa, mas evita as dificuldades de arranjos multicópia. Ela ocorre por um mecanismo que difere dos discutidos até agora, e é baseada nas propriedades de um **elemento de transposição** chamado de elemento *P*, que atua como vetor. Os elementos de transposição são o assunto do Capítulo 14.

Padrão de transmissão de T-DNA

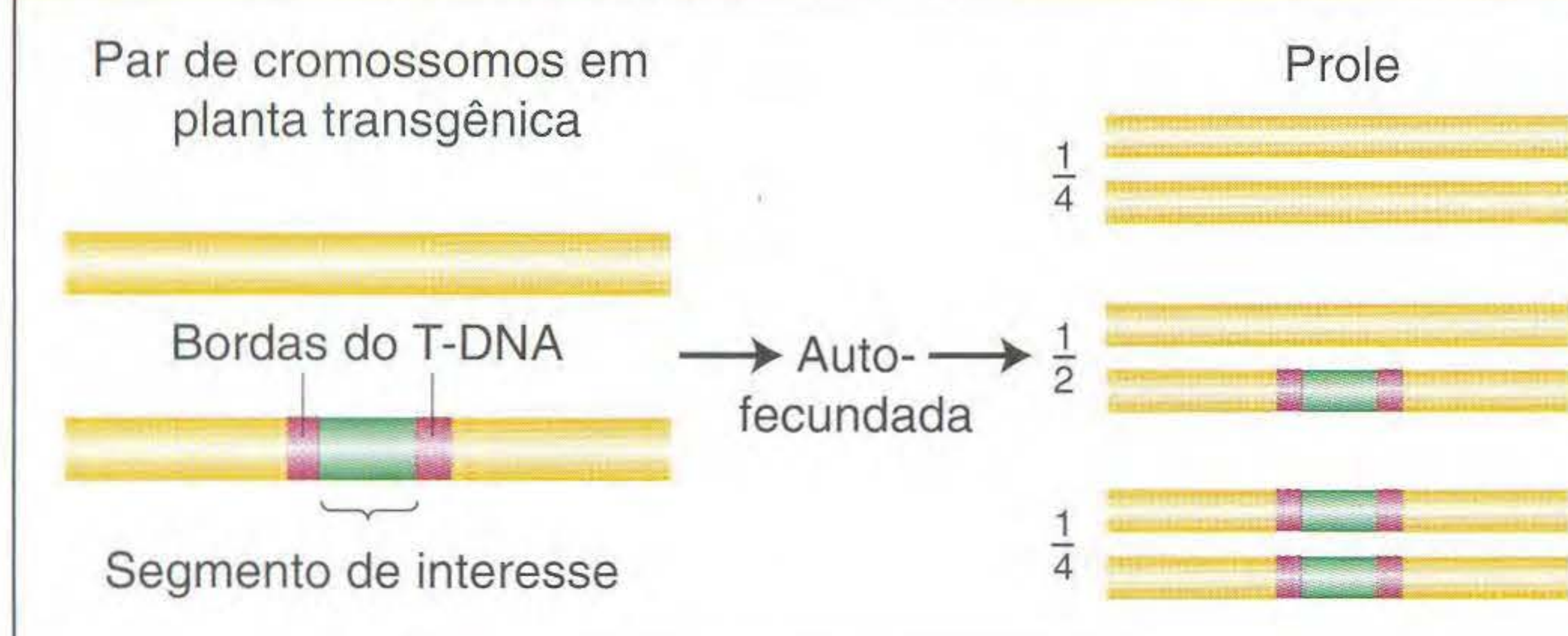


FIG. 20.27 A região de T-DNA e qualquer DNA inserido em um cromossomo de planta em uma planta transgênica são transmitidos em um padrão de herança mendeliano.

Criação de transgenes de *Caenorhabditis elegans*

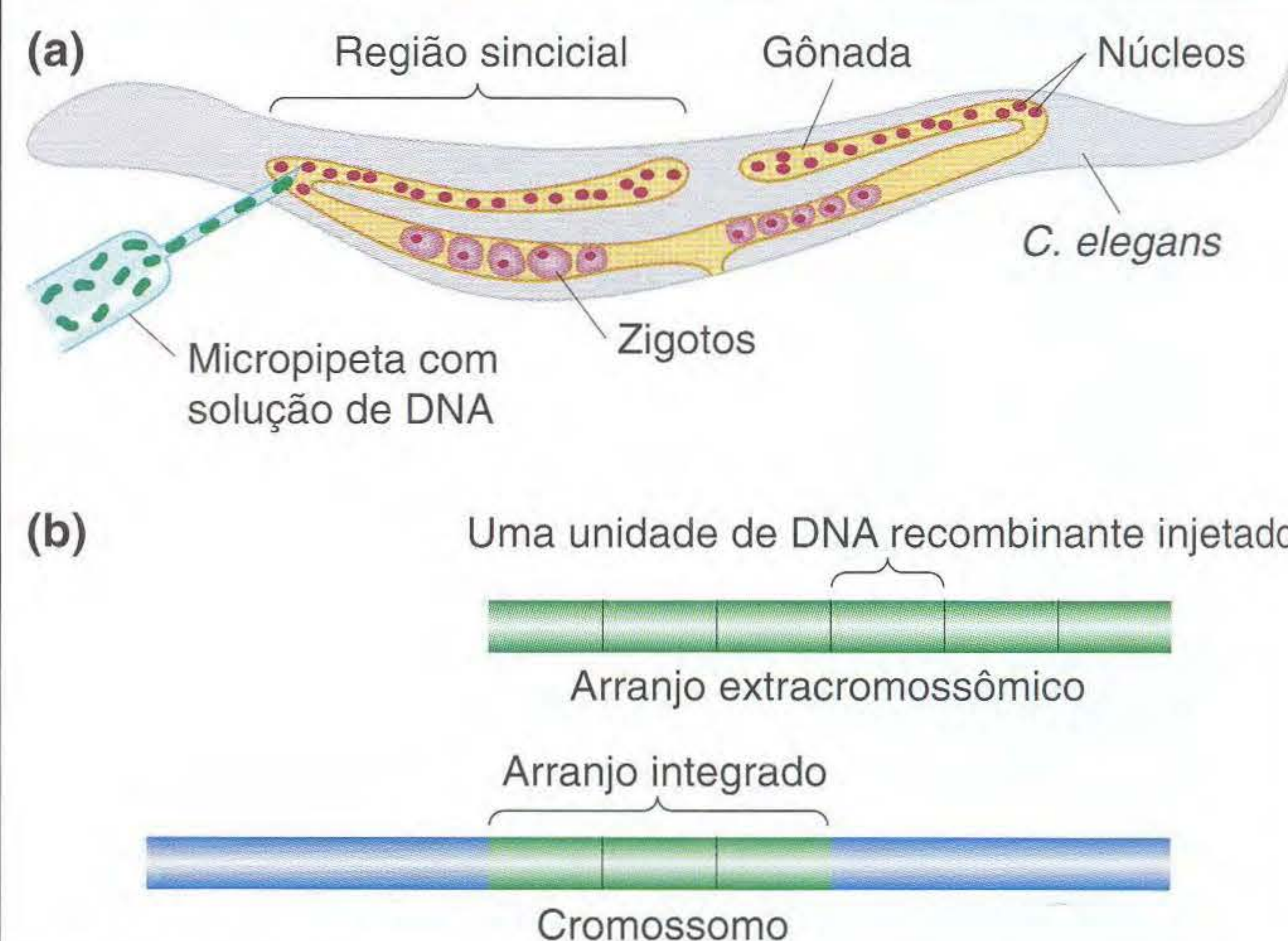


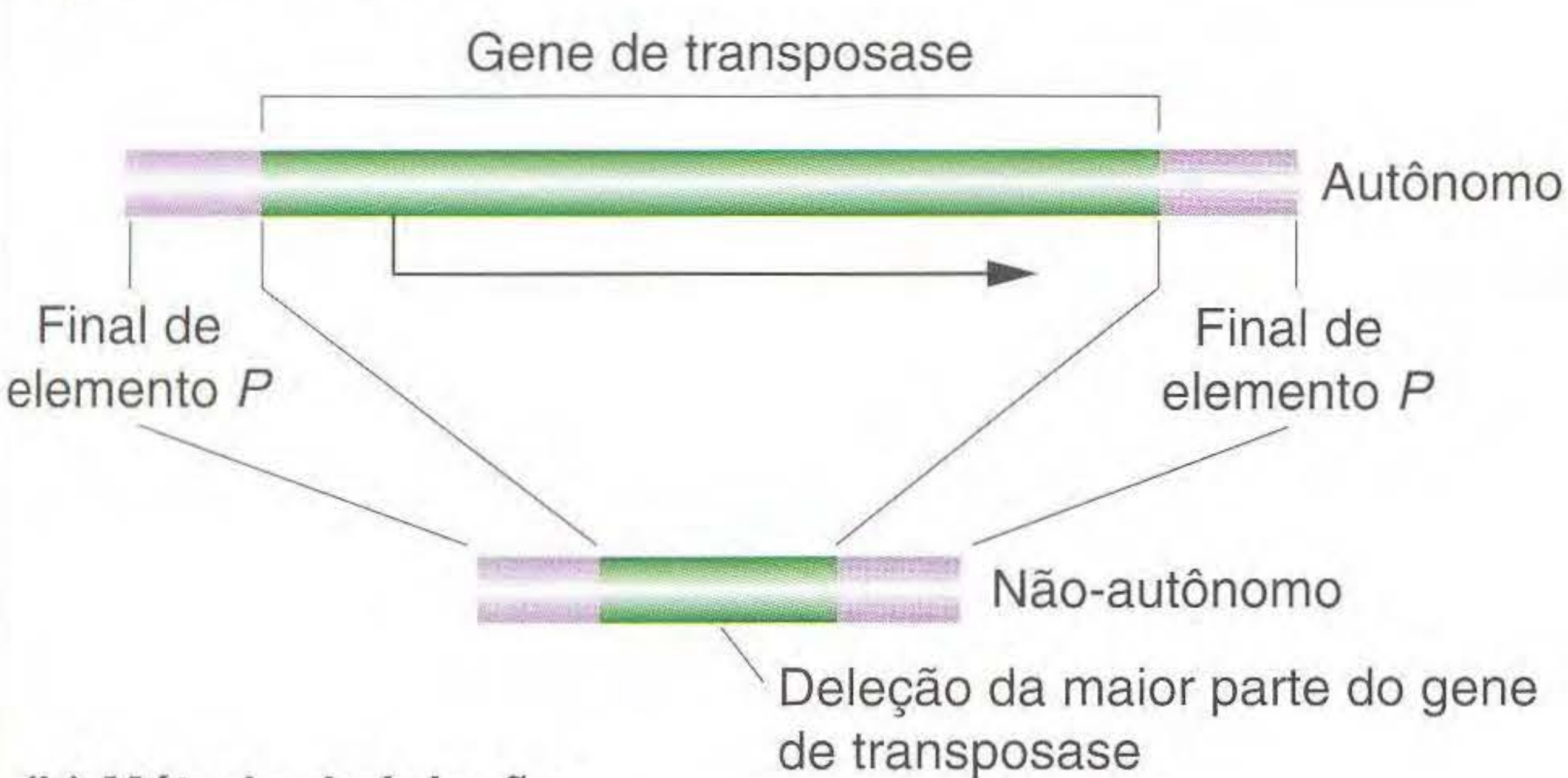
FIG. 20.28 Os transgenes de *C. elegans* são criados injetando-se DNA transgênico diretamente em uma gônada. (a) Método de injeção. (b) Os dois tipos principais de resultados transgênicos. Os arranjos extracromossômicos e arranjos integrados em locais cromossômicos ectópicos.

Para nossas finalidades aqui, tudo o que precisamos saber é que os elementos *P* existem em dois tipos (Figura 20.29a).

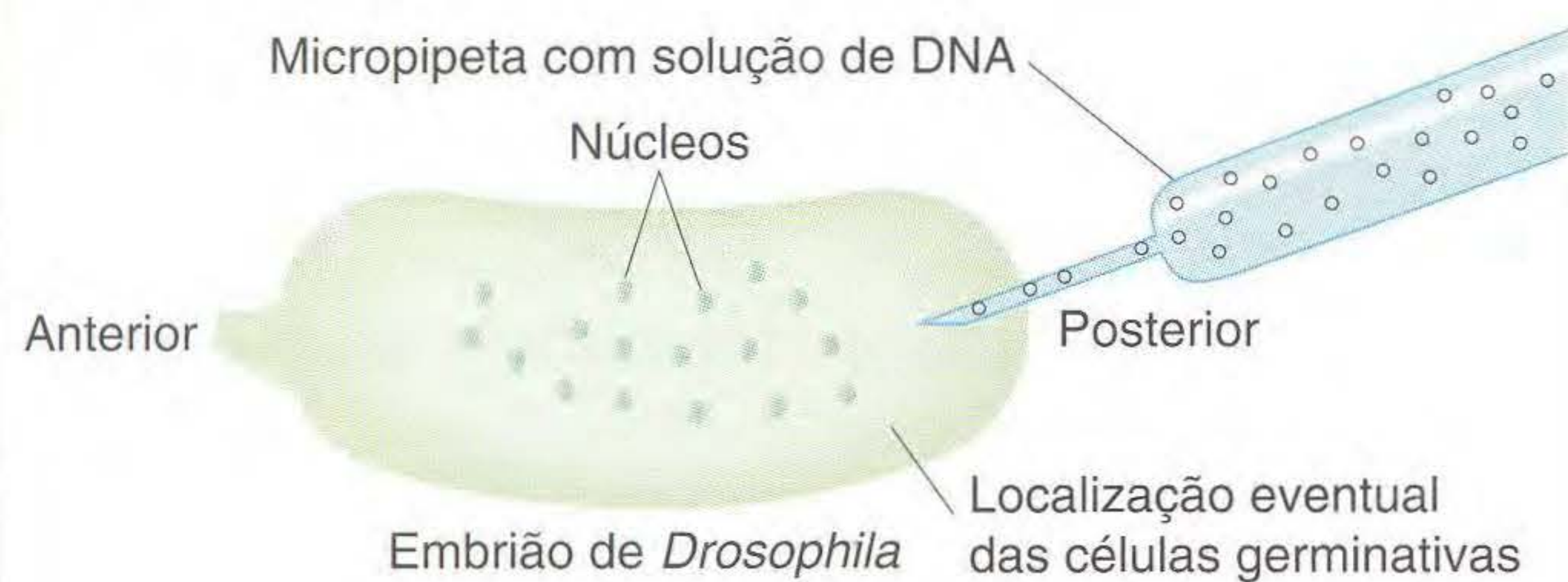
- Um tipo de elemento, com 2.912 pb de tamanho, codifica uma proteína chamada de **transposase**, que é necessária para os elementos *P* se moverem para uma nova posição no genoma. Lembre que esse tipo de elemento é chamado de “autônomo” porque pode ser transposto pela ação de sua própria enzima transposase.
- A transposase foi deletada do segundo tipo de elemento, que é, portanto, um elemento não-autônomo. Um elemento não-autônomo pode mover-se para um novo local genômico se a transposase receber um elemento autônomo. O requisito é que o elemento não-autônomo contenha os primeiros 200 pb e os 200 pb finais do elemento autônomo, que inclui seqüências que a transposase precisa para reconhecer o elemento não-autônomo. Além disso, qualquer DNA inserido entre as pontas de um elemento *P* não-autônomo será transposto também.

Criação de transgenes de *Drosophila melanogaster*

(a) Elementos *P*



(b) Método de injeção



(c) Evento integrado

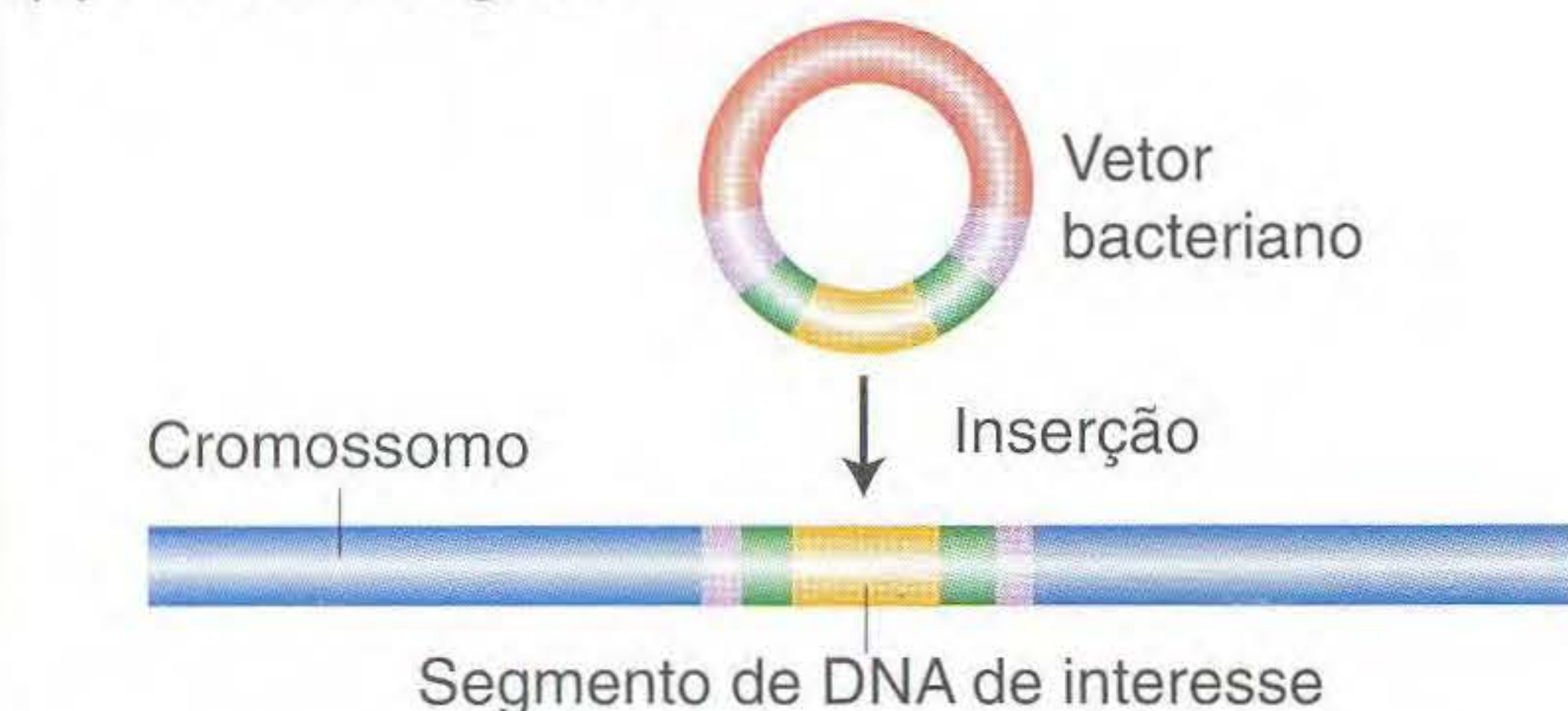


FIG. 20.29 Transgenes de *D. melanogaster* são criados injetando-se um vetor de transposição no embrião inicial de *Drosophila*. (a) A estrutura geral dos elementos *P* de transposição autônomos e não-autônomos. (b) Método de injeção. (c) O vetor circular elemento *P* (em cima) e um evento típico de integração em um local cromossômico ectópico (embaixo). Note que as seqüências de vetor bacteriano não se tornam integradas ao genoma, mas, na integração, exatamente uma cópia do segmento de DNA é contida entre as pontas do elemento *P*.

Como com *C. elegans*, o DNA é injetado em um sincício, nesse caso o embrião inicial de *Drosophila* (Figura 20.29b). Mais exatamente, o DNA é injetado no sítio de formação de células germinativas, no pólo posterior do embrião. Os adultos que crescem do embrião injetado tipicamente não expressam o transgene, mas conterão algumas células germinativas transgênicas, e essas células serão expressas na prole.

Que tipo de vetor leva o DNA injetado? Para produzir drosófilas transgênicas, precisamos injetar *dois* plasmídeos bacterianos recombinantes. Um contém o elemento *P* autônomo que fornece as seqüências codificantes para a transposase. Esse elemento é o plasmídeo *P helper*. O outro é o vetor elemento *P* levando o transgene. O vetor elemento *P* é um elemento construído não-autônomo contendo as pontas do elemento *P* e, inserido entre as pontas, o pedaço de DNA clonado que queremos incorporar como transgene no genoma da mosca. Uma solução de DNA contendo ambos esses plasmídeos é injetada no pólo posterior do embrião sincicial. A transposase *P* expressa pelo plasmídeo *P helper* injetado catalisa a inserção do vetor elemento *P* no genoma da mosca. A natureza da reação enzimática da transposase garante que apenas uma única cópia do elemento é inserida em determinado local (Figura 20.29c).

Como nós podemos detectar a prole que se desenvolve de gametas que recebem bem-sucedidamente o DNA clonado? Tipicamente, ela é detectada devido à sua expressão de um alelo transgênico tipo selvagem de um gene para o qual a linhagem receptora leva um alelo mutante recessivo.

Como mencionado no Capítulo 13, os elementos de transposição são amplamente usados na transgênese, tanto em plantas como em insetos. Talvez o exemplo de planta mais bem conhecido seja o sistema do elemento *Ac* primeiro descrito em *Zea mays* (milho), que foi desenvolvido em um vetor transgênico de clonagem para uso em muitas plantas.

Transgênese em *M. musculus* Os camundongos são os modelos mais importantes para a genética de mamíferos. Grande parte da tecnologia desenvolvida em camundongos é potencialmente aplicável em humanos. Existem duas estratégias de transgênese em camundongos, cada uma tendo suas vantagens e desvantagens:

- **Inserções ectópicas.** Os transgenes são inseridos aleatoriamente no genoma, geralmente como arranjos multicópia.
- **Gene-alvo.** A seqüência transgênica é inserida em um local ocupado por uma seqüência homóloga no genoma. Isto é, o transgene substitui sua contraparte homóloga normal.

Inserções ectópicas Para inserir transgenes em locais aleatórios, o procedimento é simplesmente injetar uma solução de DNA bactericamente clonado no núcleo de um zigoto (Figura 20.30a). Vários zigotos injetados são inseridos no oviduto feminino, onde alguns irão desenvolver-se em filhotes de camundongo. Em um estágio mais posterior, o transgene torna-se integrado aos cromossomos de núcleos aleatórios. Ocasionalmente, as células transgênicas formam parte da

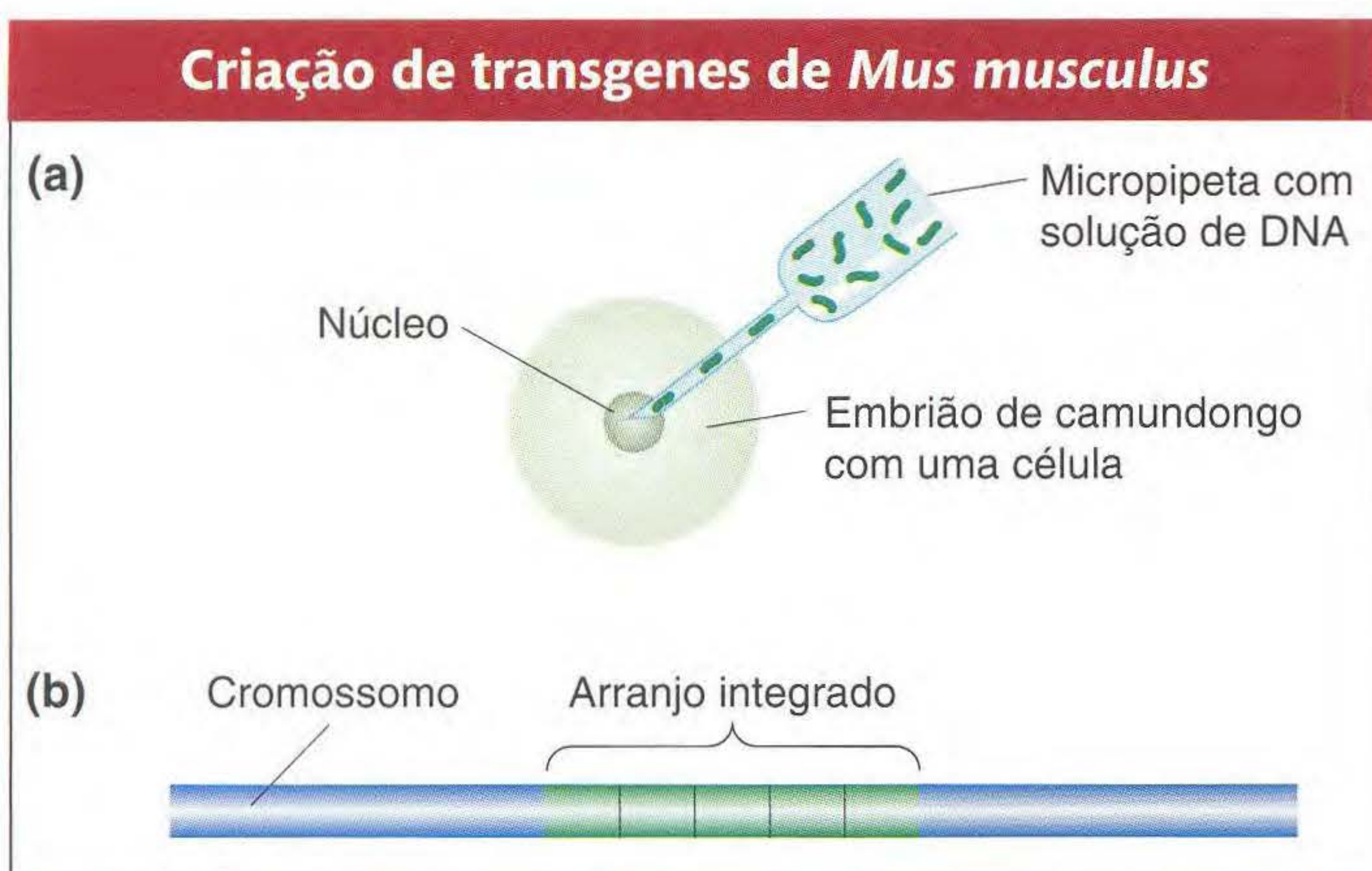


FIG. 20.30 Os transgenes de *M. musculus* são criados pela inserção em locais cromossômicos ectópicos. (a) Método de injeção. (b) Um integrante ectópico típico, com várias cópias do transgene recombinante inserido em um arranjo.

linhagem germinativa, e, nesses casos, um embrião injetado se desenvolverá em um adulto cujas células germinativas contêm o transgene inserido em alguma posição aleatória em um dos cromossomos (Figura 20.30b). Parte da prole desses adultos herdará o transgene em todas as células. Haverá um arranjo de várias cópias em cada ponto de inserção, mas o local, tamanho e estrutura dos arranjos serão diferentes para cada evento de integração. A técnica dá origem a alguns problemas: (1) o padrão de expressão dos genes inseridos aleatoriamente pode ser anormal (chamado um **efeito de posição**), porque o ambiente cromossômico local não tem as seqüências normais reguladoras do gene, e (2) podem ocorrer rearranjos de DNA dentro dos arranjos múltiplos (em essência, mutando as seqüências). Entretanto, essa técnica é muito mais eficiente e menos trabalhosa que o direcionamento do gene.

Direcionamento do gene O direcionamento do gene permite que os pesquisadores eliminem ou modifiquem as funções codificadas por um gene. Em uma aplicação, um gene mutante pode ser reparado pela **substituição gênica**, pela qual um alelo tipo selvagem substitui um mutante em sua localização cromossômica normal. A substituição gênica evita tanto o efeito de posição quanto os rearranjos de DNA associados à inserção ectópica, pois uma única cópia do gene é inserida em seu ambiente cromossômico normal.

O direcionamento do gene no camundongo é feito em células-tronco embrionárias (células ES). Em geral, uma célula-tronco é uma célula indiferenciada, em determinado tecido ou órgão, que se divide assimetricamente para produzir uma prole de células-tronco e uma célula que irá diferenciar-se em um tipo celular terminal. As células ES são células-tronco especiais que podem diferenciar-se para formar qualquer tipo celular do corpo, incluindo, sobretudo, a linhagem germinativa.

Para ilustrarmos o processo de direcionamento gênico, veremos como ela obtém um de seus resultados típicos, isto

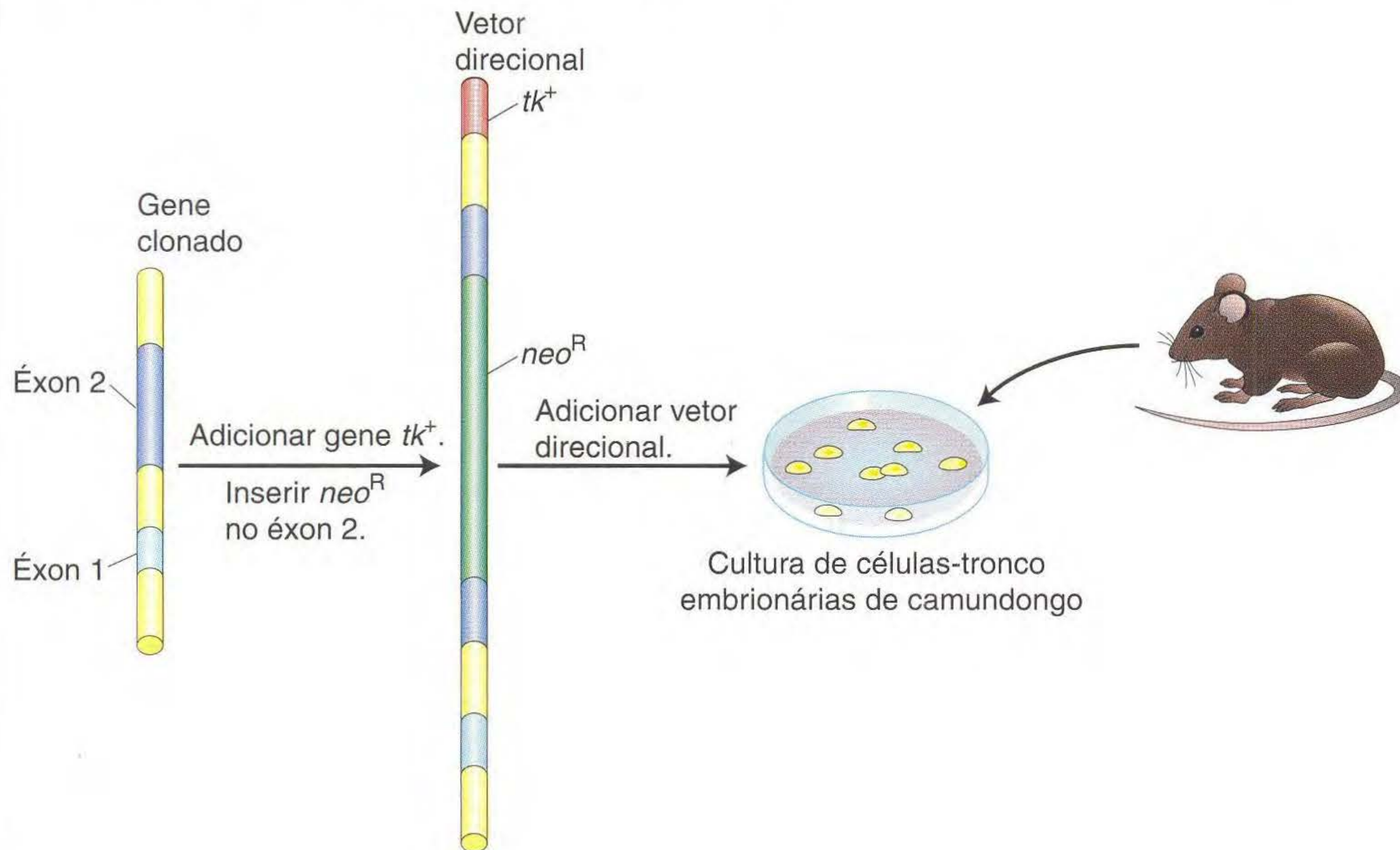
FIG. 20.31 Produção de células que contêm uma mutação em um gene específico, conhecida como mutação dirigida ou nocaute gênico. (a) Cópias de um gene clonado são alteradas *in vitro* para produzir um vetor direcional. O gene mostrado aqui foi inativado pela inserção do gene de resistência à neomicina (*neo^R*) em uma região codificante de proteína (éxon 2) do gene e foi inserido em um vetor. O gene *neo^R* servirá mais tarde como um marcador para indicar que o DNA vetor ficou residente em um cromossomo. O vetor também foi construído para levar um segundo marcador em uma ponta: o gene herpes *tk*. Esses marcadores são o padrão, mas outros também podem ser usados. Quando um vetor, com seus marcadores duplos, está completo, ele é introduzido em células isoladas do embrião de camundongo. (b) Quando ocorre recombinação homóloga (*esquerda*), as regiões homólogas do vetor, juntas com qualquer DNA no meio, mas excluindo o marcador na ponta, tomam o lugar do gene original. Esse evento é importante porque as seqüências vetores servem como um marcador útil para detecção da presença desse gene mutante. Em muitas células, entretanto, o vetor inteiro (completo com o marcador extra na ponta) insere-se ectopicamente (*meio*) ou não se integra (*embaixo*). (c) Para isolar células que levam uma mutação dirigida, todas as células são postas em um meio contendo drogas selecionadas – aqui, um análogo de neomicina (G418) e ganciclovir. G418 é letal para as células a menos que elas tenham um gene funcional *neo^R*, eliminando assim as células nas quais não ocorreu integração do DNA vetor (amarelo). Enquanto isso, o ganciclovir mata quaisquer células que abriguem o gene *tk*, eliminando assim as células portadoras de um vetor integrado aleatoriamente (vermelho). Conseqüentemente, praticamente todas as células que sobrevivem e proliferam são as que possuem a inserção dirigida (verde). [De M.R. Capecchi, "Targeted Gene Replacement." Copyright 1994 by Scientific American, Inc. Todos os direitos reservados.]

é, a substituição de um gene inativo pelo gene normal. Tal inativação dirigida é chamada de **nocaute gênico**. Primeiro, um gene clonado perturbado, que é inativo, é direcionado para substituir o gene funcional em uma cultura de células ES, produzindo células ES contendo um nocaute gênico (Figura 20.31a). As construções de DNA contendo o gene defeituoso são injetadas nos núcleos de células ES cultivadas. O gene defeituoso se insere mais frequentemente em sítios não-homólogos (ectópicos) do que em sítios homólogos (Figura 20.31b), e, assim, a etapa seguinte é selecionar as raras células nas quais o gene defeituoso foi substituído pelo gene funcional, como desejado. Como é possível selecionar células ES que contenham uma rara substituição de gene? O engenheiro genético pode incluir alelos de resistência a drogas em um DNA construído disposto de tal modo que as substituições podem ser distintas de inserções ectópicas. Um exemplo é mostrado na Figura 20.31c.

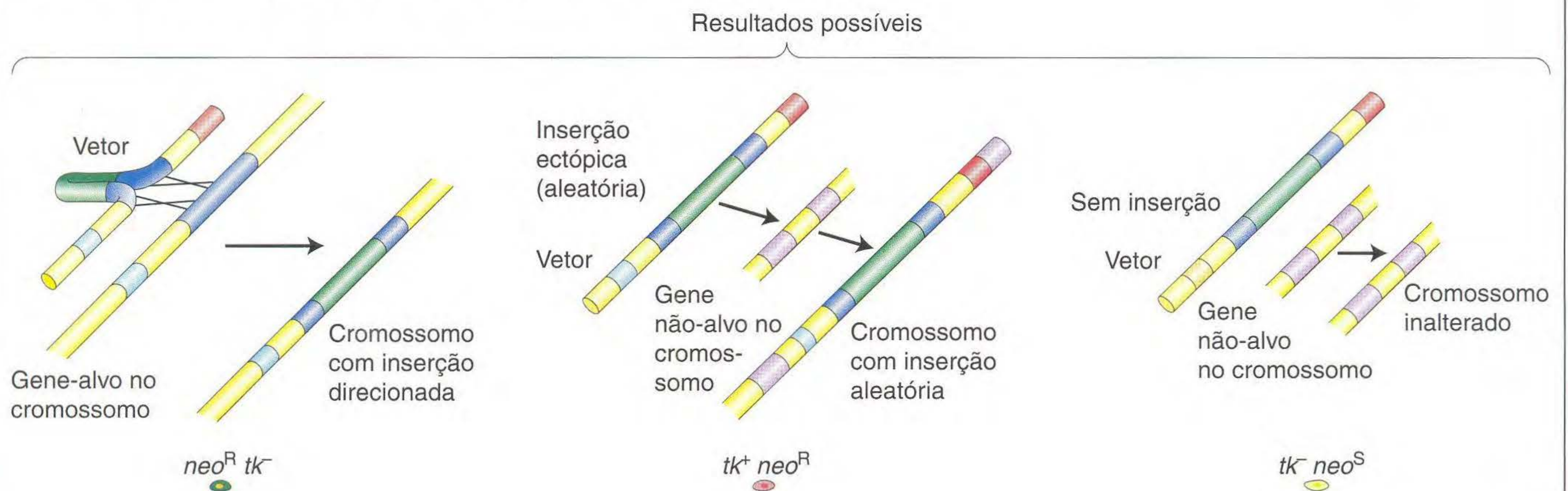
Na segunda parte do procedimento, as células ES que contêm uma cópia do gene perturbado de interesse são injetadas em um embrião inicial (Figura 20.32a). Os adultos criados desses embriões são cruzados com parceiros normais. A prole resultante é quimérica, tendo alguns tecidos derivados das linhagens originais e alguns das linhagens ES transplantadas. Os camundongos quiméricos são então cruzados com seus irmãos para produzir camundongos homozigotos com o nocaute em cada cópia do gene (Figura 20.32b). Os camundongos contendo o gene-alvo em cada uma de suas células são identificados por sondas moleculares quanto a seqüências únicas do transgene.

Células produtoras contendo um nocaute gênico direcionado

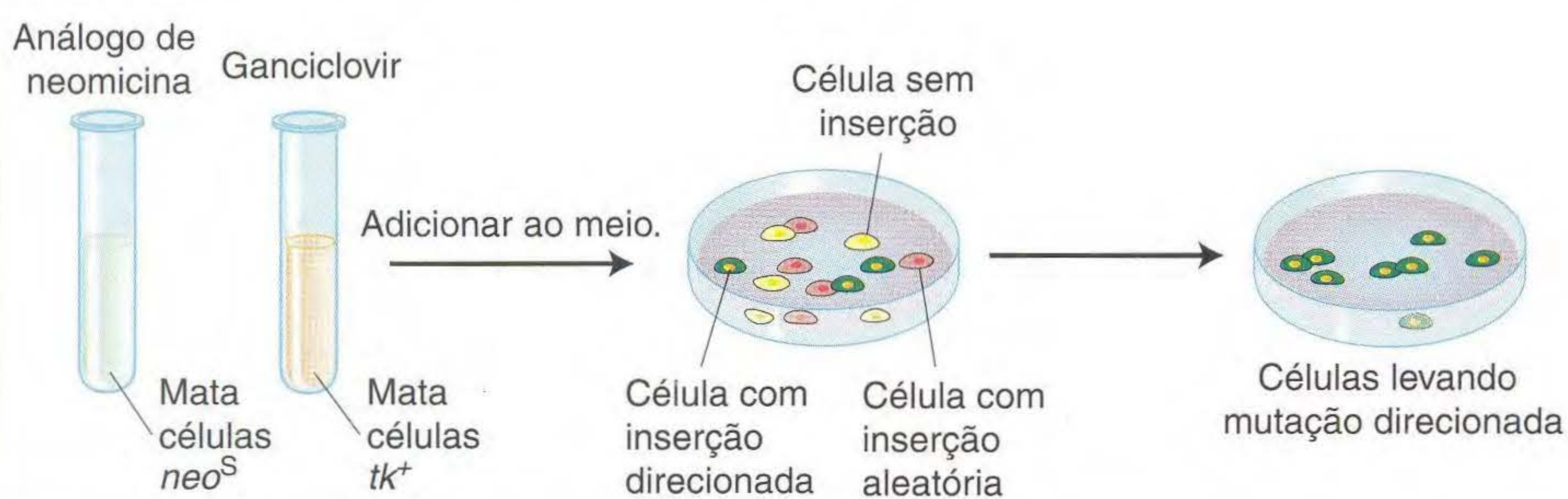
(a) Produção de células ES com um nocaute gênico



(b) Inserção direcional de DNA vetor por recombinação homóloga



(c) Seleção de células com nocaute gênico



Produção de camundongo contendo o nocaute gênico direcionado

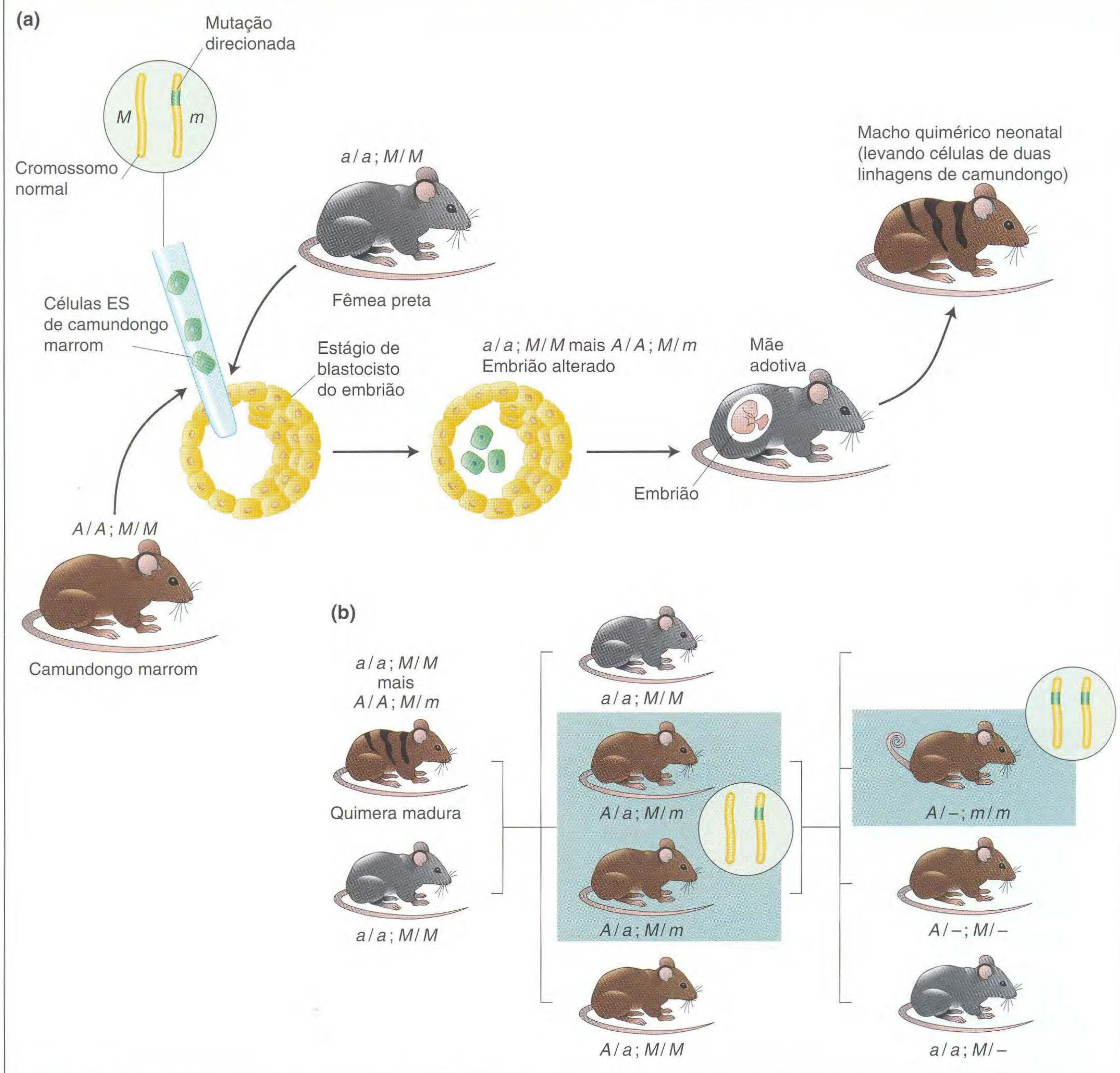


FIG. 20.32 Um camundongo nocaute é produzido pela inserção de células ES levando a mutação dirigida. (a) As células embrionárias (ES) são isoladas de um camundongo aguti (marrom) (A/A) e alteradas para levar uma mutação dirigida (m) em um cromossomo. As células ES são então inseridas nos embriões jovens, um dos quais é mostrado. A cor da pelagem dos neonatos futuros é um guia que indica se as células ES sobreviveram no embrião. Assim, as células ES são tipicamente postas em embriões que, na ausência das células ES, iriam adquirir uma pelagem totalmente preta. Tais embriões são obtidos de uma linhagem preta que não tem o alelo dominante aguti (a/a). Os embriões contendo as células ES crescem a termo em mães adotivas. Aguti misturado com preto indica os neonatos nos quais as células ES sobreviveram e proliferaram. (Tais camundongos são chamados de *quimeras* porque contêm células derivadas de duas linhagens diferentes de camundongo.) A cor totalmente preta, em contraste, indica que as células ES morreram, e esses camundongos são excluídos. A representa aguti, a preto, m é a mutação dirigida e M é seu alelo tipo selvagem. (b) Os machos quiméricos são cruzados com fêmeas pretas (não-aguti). As proles são triadas quanto à evidência da mutação dirigida (verde no *detalhe*) no gene de interesse. O exame direto dos genes no camundongo aguti revela quais desses animais (*boxes*) herdaram a mutação dirigida. Machos e fêmeas levando a mutação são cruzados para produzir camundongos cujas células levam a mutação escolhida em ambas as cópias do gene dirigido (*detalhe*) e, assim, não têm um gene funcional. Tais animais (*boxes*) são identificados definitivamente pela análise direta de seu DNA. O nocaute nesse caso resulta em um fenótipo de cauda enrolada. [De M.R. Capecchi, "Targeted Gene Replacement." Copyright 1994 by Scientific American, Inc. Todos os direitos reservados.]

As técnicas transgênicas de linhagem germinativa foram desenvolvidas para todas as espécies eucarióticas bem estudadas. Essas técnicas dependem de uma compreensão da biologia reprodutiva da espécie receptora.

Terapia gênica humana

A meta geral da terapia gênica é atacar a base genética da doença em sua fonte: “curar” ou corrigir uma condição anormal causada por um alelo mutante introduzindo um alelo tipo selvagem nas células. Essa técnica foi bem-sucedidamente aplicada em muitos organismos experimentais, e tem o potencial, em humanos, de corrigir algumas doenças hereditárias, particularmente as associadas a diferenças monogênicas. Embora a terapia gênica tenha sido tentada para várias dessas doenças, até agora não existem relatos claros de

sucesso. Entretanto, as implicações da terapia gênica são tão amplas que o enfoque merece ser aqui considerado.

Para compreender o enfoque, considere um exemplo mostrando como a terapia gênica corrigiu uma deficiência de hormônio de crescimento (Figura 20.33). Os camundongos com a mutação recessiva *little* (*lit*) são anões porque não têm uma proteína (o receptor hormonal liberador do hormônio de crescimento, ou GHRHR) necessária para induzir a hipófise a secretar o hormônio de crescimento de camundongo no sistema circulatório. A etapa inicial para corrigir essa deficiência foi injetar cerca de 5.000 cópias de um transgene em zigotos homozigotos *lit/lit*. Essa construção era um fragmento de DNA linear de 5 kb que continha seqüências codificantes de hormônio de crescimento de rato (*RGH*) fusionadas a seqüências reguladoras para o gene de metalotioneína de camundongo. Essas seqüências reguladoras levam à expressão de qualquer gene imediatamente adjacente na presen-

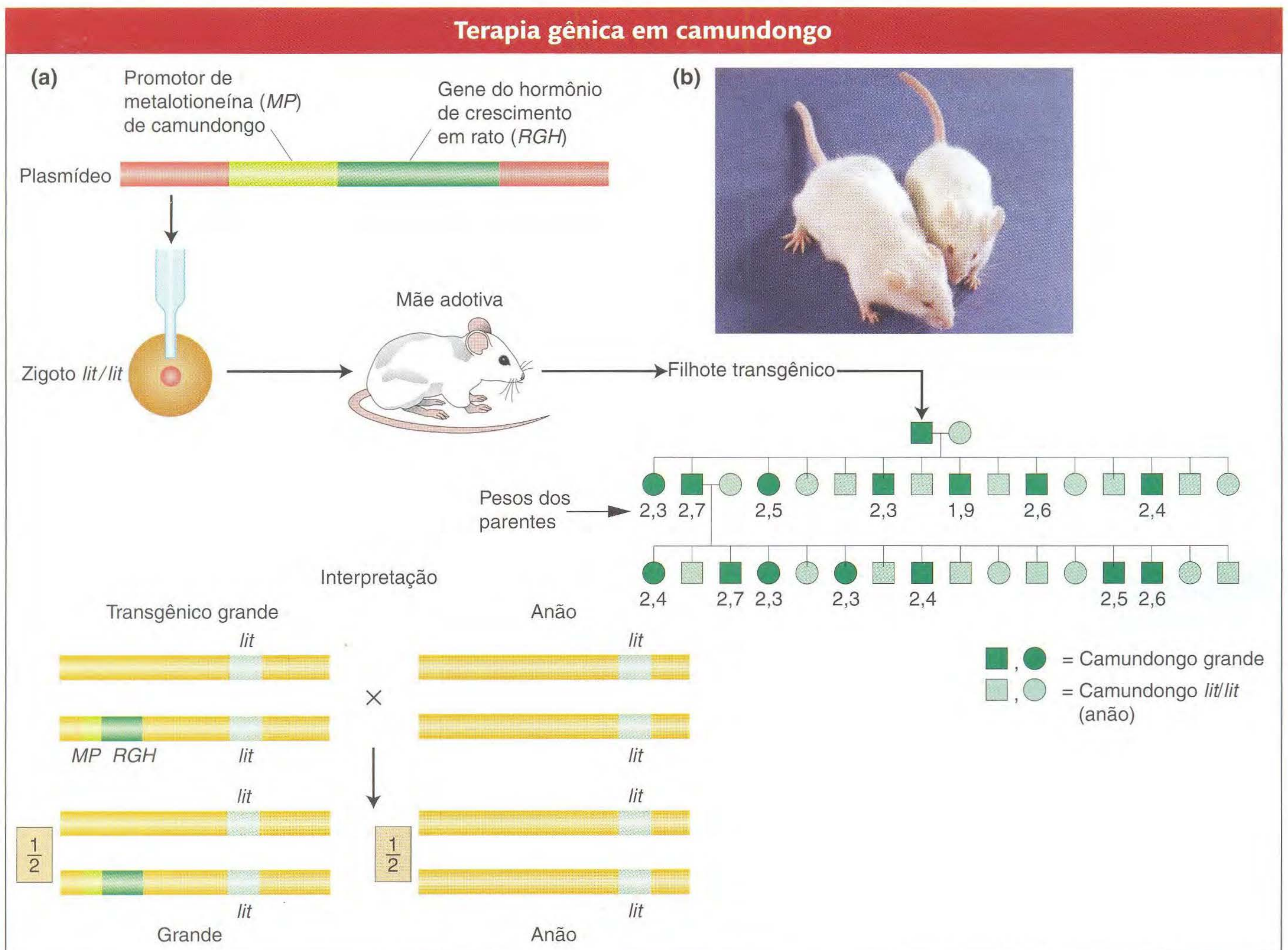


FIG. 20.33 Produção de um camundongo levando um transgene terapêutico para corrigir uma deficiência de hormônio de crescimento. (a) O gene do hormônio de crescimento de rato (*RGH*), sob o controle de um promotor de camundongo que responde a metais pesados, é inserido em um plasmídeo e usado para produzir um camundongo transgênico. *RGH* compensa o inerente nanismo (*lit/lit*) do camundongo. *RGH* é herdado de modo mendeliano dominante no heredograma anexo do camundongo. (b) Camundongo transgênico. Os camundongos são irmãos, mas o camundongo da esquerda foi derivado de um zigoto transformado pela injeção de um novo gene composto do promotor de metalotioneína de camundongo fusionado ao gene estrutural do hormônio de crescimento de rato. (Esse camundongo pesa 44 g, e seu irmão não-tratado pesa 29 g.) O novo gene é transmitido para a prole de modo mendeliano, e, assim, é provado como estando cromossomicamente integrado. [De R.L. Brinster.]

ça de metais pesados. Os zigotos foram então implantados nos úteros de camundongos adotivos, e os filhotes nasceram e foram criados. Cerca de 1% deles era transgênico, apresentando tamanho grande quando metais pesados eram administrados durante o desenvolvimento. Um camundongo transgênico representativo foi então cruzado com uma fêmea homozigota *lit/lit*. O pedigree seguinte é mostrado na Figura 20.33a. Aqui, vemos que os camundongos com duas a três vezes o peso de seus parentes *lit/lit* são produzidos em gerações subsequentes (Figura 20.33b). Esses camundongos maiores são sempre heterozigotos nesse pedigree, mostrando que o transgene do hormônio de crescimento de rato atua como um alelo dominante. Assim, a introdução do transgene *RGH* obteve uma “cura” no sentido de que a prole não apresentou o fenótipo anormal.

Esse exemplo particular mostra alguns pontos importantes sobre o processo de terapia gênica. O defeito genético ocorre no gene codificante de GHRHR, um regulador da produção de hormônio de crescimento. Entretanto, a terapia gênica não é uma tentativa de corrigir o defeito original no gene codificante GHRHR. A terapia gênica ultrapassa a necessidade de GHRHR e produz hormônio por outra via, especificamente pela introdução de um gene de hormônio de crescimento de rato que não está sob o controle de GHRGH. Esse transgene é expresso sob o controle de um promotor induzível e nos tecidos onde GHRHR não é necessário para a liberação do hormônio de crescimento. (Você pode indagar por que o hormônio de crescimento de rato foi usado em vez do hormônio de crescimento de camundongo. O gene recombinante do hormônio de crescimento de rato produz tanto mRNA quanto proteína com seqüências distinguíveis das versões de camundongo, e ambas as moléculas podiam ser diretamente dosadas.)

Passemos agora à condição dos vários enfoques técnicos. Podem ser aplicados dois tipos básicos de terapia gênica aos

humanos: linhagens germinativa e somática. A meta da **terapia gênica de linhagem germinativa** (Figura 20.34a) é a mais ambiciosa: introduzir células transgênicas na linhagem germinativa bem como na população de células somáticas. Não só esse tipo de terapia obtém a cura da pessoa tratada, mas seus filhos também levariam o transgene terapêutico. A cura do defeito recessivo do camundongo *lit* é um exemplo de terapia gênica de linhagem germinativa. No momento, essas tecnologias dependem da integração ectópica ou substituição de gene que ocorre aleatoriamente, e esses eventos são suficientemente raros, tornando a terapia gênica de linhagem germinativa impraticável por enquanto.

Terapia gênica somática (Figura 20.34b) Tenta corrigir um fenótipo de doença tratando *algumas* células somáticas na pessoa afetada. Nenhum transgene atinge a linhagem germinativa. No momento, não é possível tornar todo o corpo transgênico, e, assim, o método aborda doenças causadas por genes que são predominantemente expressos em um tecido. Em tais casos, é provável que nem todas as células desse tecido precisem se tornar transgênicas. Uma parte das células portadoras do transgene pode aliviar os sintomas gerais da doença. O método continua removendo algumas células de um paciente com o genótipo defeituoso, e tornando essas células transgênicas pela introdução de cópias do gene tipo selvagem clonado. As células transgênicas são então reintroduzidas no corpo do paciente, onde fornecem o funcionamento gênico normal.

Os vírus foram usados como vetores para levar o DNA para os genomas de muitos animais, inclusive humanos. No Capítulo 14, consideramos o uso de um retrovírus contendo o gene normal codificante de adenosina desaminase para tratar a doença da imunodeficiência combinada grave em crianças. Lembre que alguns pacientes com SCID desenvolveram leucemia após a terapia gênica, possivelmente como

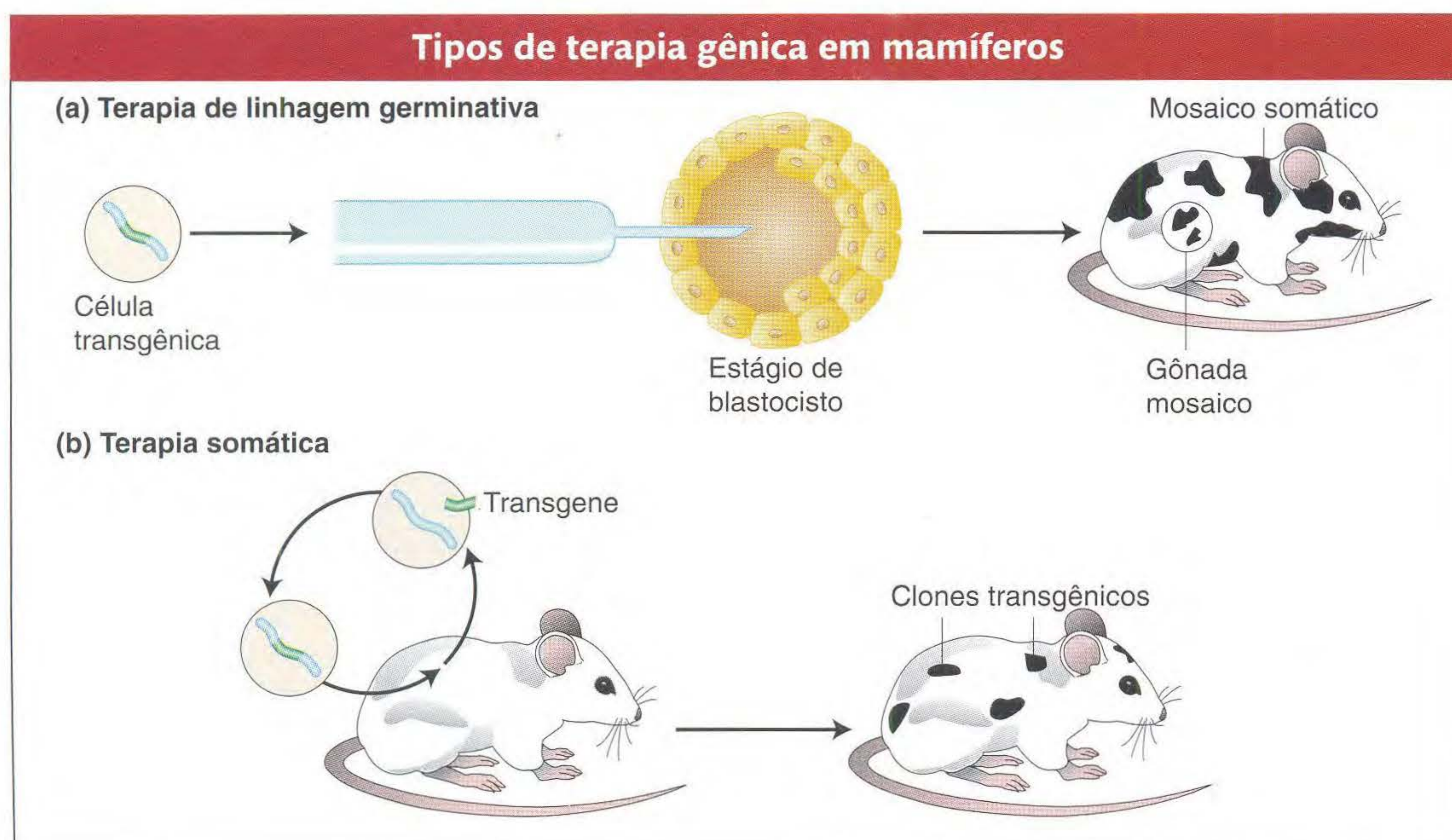


FIG. 20.34 A terapia com linhagem germinativa introduz um transgene na linhagem germinativa, bem como células somáticas, de modo que ele possa ser herdado. A terapia de células somáticas introduz o transgene em algumas das células somáticas.

resultado de inativação gênica. Outro problema é que um retrovírus infecta apenas células proliferativas, tais como células sangüíneas, não podendo assim ser usado para tratar muitas doenças hereditárias que afetam tecidos nos quais as células raramente ou nunca se dividem.

Outro vetor usado na terapia gênica humana é o adenovírus. Esse vírus normalmente infecta epitélios respiratórios, injetando seu genoma em células epiteliais que revestem a superfície do pulmão. O genoma viral não se integra ao cromossomo, mas persiste extracromossomicamente nas células, o que elimina o problema do vetor inativando um gene residente. Outra vantagem do adenovírus como um vetor é que ele ataca células que não se dividem, e, assim, a maioria dos tecidos são em princípio suscetíveis. Os adenovírus são uma escolha apropriada do vetor para tratar a fibrose cística, uma doença do epitélio respiratório. A terapia gênica para a fibrose cística está sendo tentada introduzindo vírus que possuem o alelo tipo selvagem de fibrose cística por via nasal com um *spray*.

Embora exista algum motivo para acreditar que as dificuldades técnicas da terapia gênica somática serão superadas, essas dificuldades são consideráveis. Uma delas é como direcionar o sistema de inserção do transgene para o tecido apropriado para determinada doença. Outra é como construir o transgene para assegurar consistentemente altos níveis de expressão. Ainda outra é como proteger contra efeitos colaterais potencialmente prejudiciais, tais como os que podem ser causados pela má expressão do gene transgênico. Elas são áreas importantes de pesquisa da terapia gênica.

As tecnologias de transgênese estão atualmente sendo aplicadas a humanos com a meta específica de aplicação da correção de certos distúrbios herdáveis. Os desafios técnicos, sociais e éticos dessas tecnologias são consideráveis, e são áreas ativas de pesquisa e debate.

► Resumo

O DNA recombinante é feito cortando o DNA doador em fragmentos que são colocados em vetores individuais de DNA. O DNA vetor geralmente é um plasmídeo bacteriano ou DNA viral. O DNA doador e o DNA vetor são cortados pela mesma endonuclease de restrição em seqüências específicas. O DNA vetor e o doador são unidos em um tubo de ensaio para formar ligações fosfodiéster covalentes, fazendo um arcabouço fosfato-açúcar intacto para cada filamento de DNA.

A construção DNA doador-vetor é amplificada dentro das células hospedeiras enganando a maquinaria básica de replicação da célula em moléculas recombinantes replicantes. Assim, o vetor deve conter todos os sinais necessários para a replicação apropriada e segregação nessa célula hospedeira. Para sistemas baseados em plasmídeos, o vetor deve incluir uma origem de replicação e marcadores selecionáveis, tais como resistência a drogas que podem ser usadas para garantir que o plasmídeo não é perdido da célula hospedeira. No bacteriófago, o vetor deve incluir todas as seqüências necessárias para levar o bacteriófago (e o DNA “carona” exógeno) através do ciclo de crescimento lítico. O resultado da amplificação são as múltiplas cópias de cada DNA recombinante, chamadas de clones.

Geralmente, encontrar um clone específico requer uma triagem de toda uma biblioteca genômica. Uma biblioteca genômica é um conjunto de clones, embalados no mesmo vetor, que, juntos, representam todas as regiões do genoma do organismo em questão. O número de clones que constituem uma biblioteca genômica depende (1) do tamanho do genoma em questão e (2) do tamanho do inserto tolerado pelo sistema particular de vetor de clonagem. Similarmente, uma biblioteca de cDNA é uma representação do mRNA total produzido por determinado tecido ou estágio desenvolvimental em determinado organismo. Uma comparação de uma região genômica e seu cDNA pode ser uma fonte de *insight* para os locais de início e fim da transcrição e limites entre íntrons e éxons.

As sondas de DNA ou RNA unifilamentares marcadas são “iscas” importantes para pescar seqüências similares ou idênticas de misturas complexas de moléculas, seja nas bibliotecas genômicas ou de cDNA, seja nas transferências de Southern ou Northern. O princípio geral da técnica para identificar clones ou fragmentos de gel é criar uma “imagem” das colônias ou placas em uma cultura de placa de Petri, ou de ácidos nucleicos que foram separados em um campo elétrico passado por uma matriz de gel. O DNA ou RNA é então desnaturado e fixado com uma sonda desnaturada, marcada com um corante fluorescente ou uma sonda radioativa. Após a sonda não-ligada ter sido removida, o local da sonda é detectado seja pela observação de sua fluorescência, seja, se radioativa, pela exposição da amostra a um filme de raios X. As localizações da sonda correspondem às localizações do DNA ou RNA relevante na placa de Petri original ou gel de eletroforese.

A reação em cadeia da polimerase é um método poderoso para a amplificação direta de uma seqüência relativamente pequena de DNA de dentro de uma mistura complexa de DNA, sem a necessidade de uma célula hospedeira ou de muito material inicial. A chave é ter primers que são complementares a regiões flangeadoras a cada um dos dois filamentos de DNA. Essas regiões atuam como sítios de polimerização. Várias rodadas de desnaturação, priming (inclusão dos primers) e polimerização amplificam a seqüência de interesse exponencialmente.

Os grandes recursos genômicos estão tornando cada vez mais possível isolar genes apenas pelo conhecimento de sua posição em um mapa genético. O procedimento geral é uma estratégia genética direta chamada de clonagem posicional. Tanto mutantes quanto variantes naturais foram usados como pontos iniciais na descoberta do gene. Com o seqüenciamento do genoma humano e a disponibilidade de famílias com distúrbios herdados, as estratégias de clonagem posicio-

nal levaram ao isolamento de genes que, quando mutados, produzem doenças humanas, tais como a fibrose cística. As estratégias de clonagem posicional também foram usadas para isolar muitos outros genes, como, por exemplo, os genes selecionados durante a domesticação de cultivos de plantas tais como o milho.

Os transgenes são moléculas de DNA construídas que são introduzidas e expressas em células eucarióticas. Eles podem ser usados para construir uma nova mutação ou para estudar as seqüências reguladoras que constituem parte de um gene. Os transgenes podem ser introduzidos como moléculas extracromossômicas ou podem ser integrados a um cromossomo, seja

em locais aleatórios (ectópicos), seja em lugar do gene homólogo, dependendo do sistema. Tipicamente, os mecanismos usados para introduzir um transgene dependem de uma compreensão e exploração da biologia reprodutiva do organismo.

A terapia gênica é a extensão da tecnologia transgênica ao tratamento de doenças humanas. Para corrigir as condições da doença, as tentativas de terapia de células somáticas tentam introduzir em tecidos somáticos específicos um transgene que ou substitui o alelo mutante, ou suprime o fenótipo mutante. A terapia da linhagem germinativa tenta introduzir um transgene na linhagem germinativa que ou corrige o defeito mutante ou o contorna.

► Termos-chave

amniocentese	DNA ligase	reação em cadeia da polimerase (PCR)
andar no cromossomo	DNA recombinante	recuperação de mutante
anticorpo	efeito de posição	(complementação funcional)
auto-radiograma	elemento de transposição	seqüenciamento de Sanger (didesoxi)
biblioteca de cDNA	eletroforese em gel	seqüenciamento didesoxi (de Sanger)
biblioteca genômica	engenharia genética	sonda
clonagem de DNA	enzima de restrição	substituição gênica
clonagem posicional	fragmento de restrição	tecnologia do DNA
complementação funcional	genética direta	tendenciosidade de códon
(recuperação de mutante)	genômica	terapia gênica de linhagem
cosmídeo	hibridização	germinativa
cromossomo artificial bacteriano	nocautê gênico	terapia gênica somática
(BAC)	organismo geneticamente modificado	transferência de Southern
cromossomo artificial baseado em P1	(GMO)	transferência Northern
(PAC)	organismo transgênico	transgene
cromossomo artificial de levedura	palíndromo de DNA	transposase
(YAC)	plasmídeo Ti	vetor
DNA complementar (cDNA)	punção de vilosidades coriônicas	
DNA doador	(CVS)	

► Problemas Resolvidos

Problema resolvido 1. No Capítulo 9, estudamos a estrutura das moléculas de tRNA. Suponha que você queira clonar um gene de fungo que codifica um certo tRNA. Você tem uma amostra do tRNA purificado e um plasmídeo de *E. coli* que contém um único sítio de corte de *EcoRI* em um gene *tet^R* (resistência à tetraciclina), bem como um gene para resistência à ampicilina (*amp^R*). Como você clona o gene de interesse?

SOLUÇÃO

Você pode usar o próprio tRNA ou uma cópia de cDNA clonado dele para sondar o DNA contendo o gene. Um método é digerir o DNA genômico com *EcoRI* e então misturá-lo com o plasmídeo, que você também cortou com *EcoRI*. Após a transformação de uma receptora *amp^S tet^S*, selecione colô-

nias *Amp^R*, indicando transformação bem-sucedida. Das colônias *Amp^R*, selecione as colônias que são *Tet^S*. Essas colônias *Tet^S* conterão vetores com insertos no gene *tet^R*, e um grande número delas é necessário para fazer uma biblioteca. Teste a biblioteca usando o tRNA como sonda. Esses clones que se hibridizam com a sonda conterão o gene de interesse.

Alternativamente, você pode submeter DNA genômico digerido com *EcoRI* a uma eletroforese em gel e, então, identificar a banda correta sondando com o tRNA. Essa região do gel pode ser removida e usada como uma fonte de DNA enriquecido para clonar no plasmídeo cortado com *EcoRI*. Você então sonda esses clones com o tRNA para confirmar que eles contêm o gene de interesse.

► Problemas

PROBLEMAS BÁSICOS

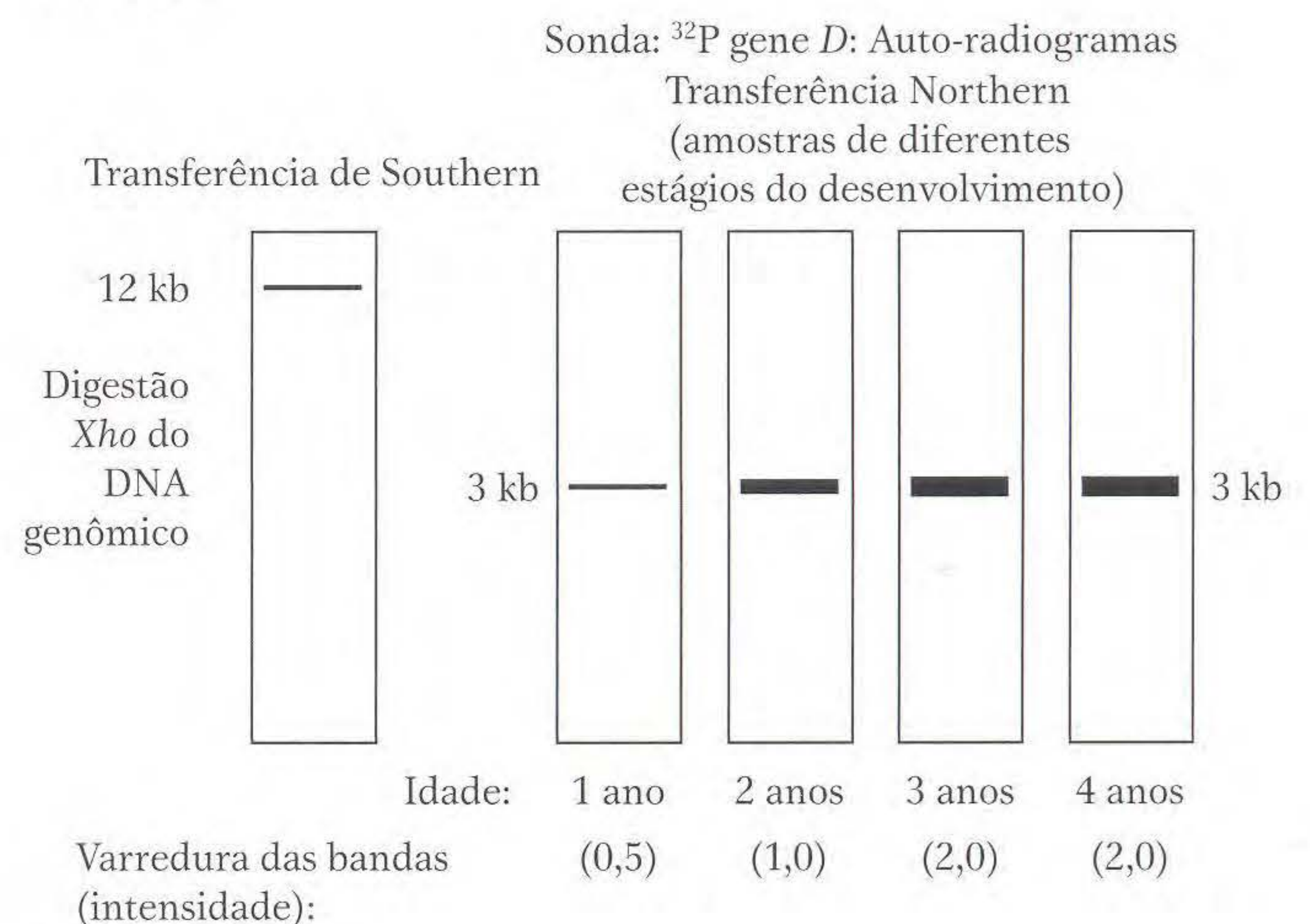
1. Deste capítulo, faça uma lista de todos os exemplos de (a) hibridização de DNA unifilamentares e (b) proteínas que se ligam ao DNA e então atuam nele.
2. Qual o uso do hidróxido de sódio para fazer cDNA?
3. Compare e contraste o uso da palavra *recombinante* como aparece nas frases (a) DNA recombinante e (b) frequência de recombinantes.
4. Por que a ligase é necessária para fazer DNA recombinante? Qual seria a consequência imediata do processo de clonagem se alguém esquecesse de adicioná-la?
5. No processo de PCR, se for suposto que cada ciclo demora 5 minutos, quantas vezes a amplificação seria feita em 1 hora?
6. A posição do gene para a proteína actina no fungo haplóide *Neurospora* é conhecida pela sequência completa do genoma. Se você tem um mutante de crescimento lento que suspeita ser um mutante de actina e quer verificar se realmente é, você (a) clonaria o mutante usando sítios de restrição convenientes flanqueando o gene de actina e então o seqüenciando, ou (b) amplificaria a sequência mutante usando PCR e então o seqüenciando?
7. Você obtém a sequência de DNA de um mutante de um gene de 2 kb no qual você está interessado, e ele apresenta diferenças de bases em três posições, todas em códons diferentes. Uma é uma mutação silenciosa, mas as outras duas são mudanças de sentido trocado (codificam novos aminoácidos). Como você demonstra que essas mudanças são mutações reais e não erros de seqüenciamento? (Suponha o seqüenciamento 99,9% preciso.)
8. Em uma transformação de T-DNA de uma planta com um transgene de um fungo (não encontrado nas plantas), a suposta planta transgênica não expressa o fenótipo esperado do transgene. Como você demonstra que o transgene está de fato presente?
9. Como você produz um camundongo que é homozigoto para um transgene de hormônio de crescimento de rato?
10. Por que foi usado cDNA e não DNA genômico na clonagem comercial do gene de insulina humana?
11. Após o DNA de *Drosophila* ter sido tratado com uma enzima de restrição, os fragmentos são inseridos em plasmídeos e selecionados como clones em *E. coli*. Com o uso dessa técnica "shotgun", cada sequência de DNA de drosófila em uma biblioteca pode ser recuperada.
 - a. Como você identificaria um clone que contém DNA codificante da proteína actina, cuja sequência de aminoácidos é conhecida?
 - b. Como você identificaria um clone codificante de um tRNA específico?
12. Em qualquer célula eucariótica transformada (digamos, de *Neurospora*), como você poderia dizer se o DNA transformante (levado em um vetor bacteriano circular)
 - a. substituiu o gene residente da receptora por duplo crossing over ou por crossing simples?
 - b. foi inserido ectopicamente?

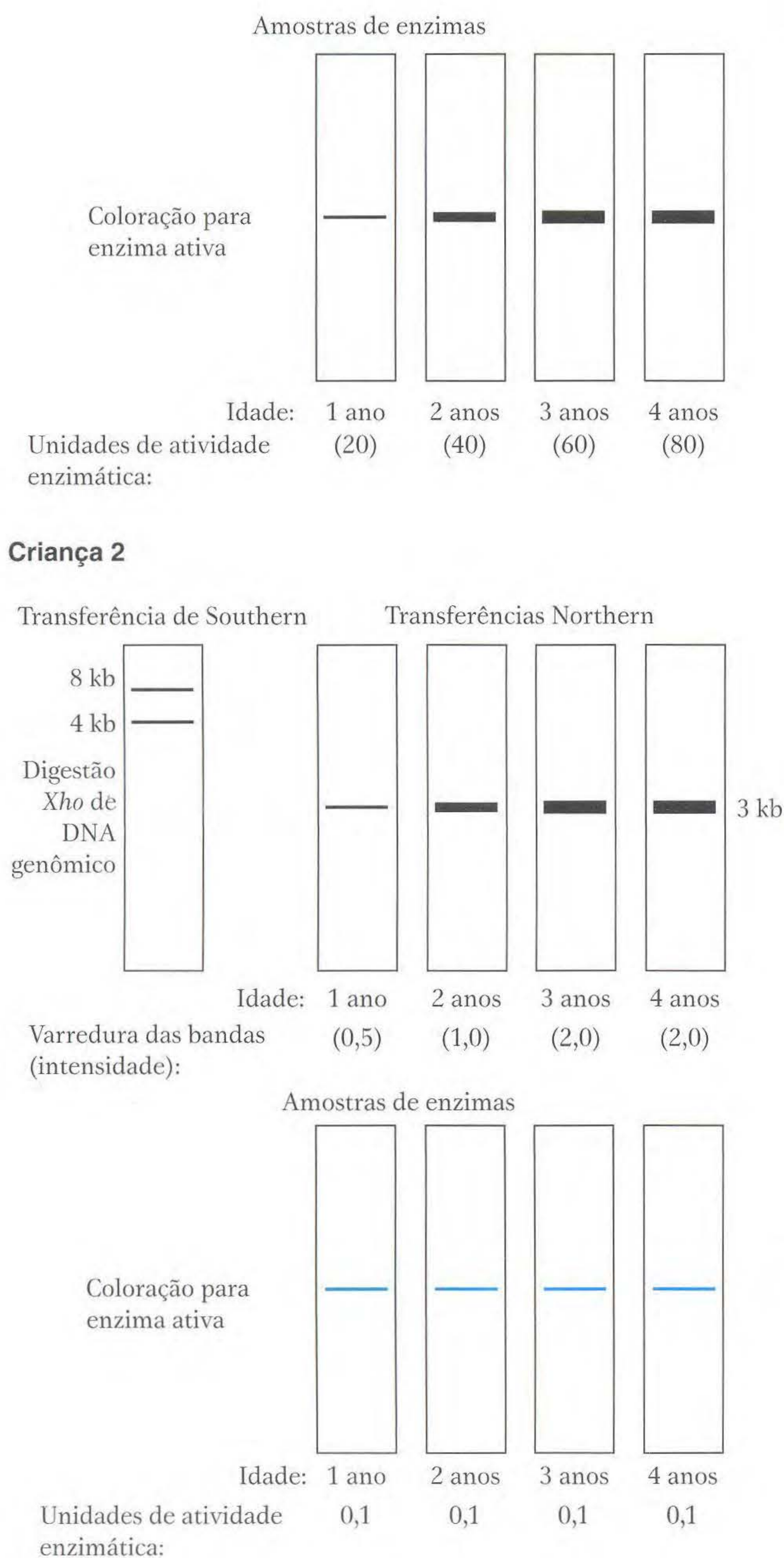
13. Em um gel de eletroforese através do qual é aplicado um pulso de campo elétrico alternado, o DNA do fungo haplóide *Neurospora crassa* ($n = 7$) move-se lentamente, mas acaba formando sete bandas, que representam frações do DNA de tamanhos diferentes e, portanto, moveram-se em velocidades diferentes. Essas bandas são supostamente os sete cromossomos. Como você mostraria qual banda corresponde a qual cromossomo?
14. A proteína codificada pelo gene de alcaptonúria tem 445 aminoácidos de tamanho, embora o gene tenha 60 kb. Como é possível essa diferença?
15. Em leveduras, você seqüenciou um pedaço de DNA tipo selvagem e ele claramente contém um gene, mas você não sabe que gene é. Portanto, para investigar mais, você gostaria de encontrar seu fenótipo mutante. Como você usaria o gene tipo selvagem clonado para fazer isso? Mostre claramente as etapas de seu experimento.

PROBLEMAS DESAFIADORES

16. A prototrofia geralmente é o fenótipo selecionado para detectar transformantes. As células prototróficas são usadas para extração do DNA do doador; então, esse DNA é clonado e os clones são adicionados a uma cultura receptora auxotrófica. Transformantes bem-sucedidos são identificados plaqueando a cultura receptora em meio mínimo e procurando colônias. Que teste experimental você usaria para ter certeza de que uma colônia que você deseja que seja transformante de fato não é:
 - a. uma célula prototrófica que entrou na cultura recipiente como um contaminante?
 - b. um revertente (mutação de volta à prototrofia por uma segunda mutação no gene mutado original) da mutação auxotrófica?
17. Duas crianças são investigadas quanto à expressão de um gene (*D*) que codifica uma enzima importante para o desenvolvimento do músculo. Os resultados dos estudos do gene e seu produto são os seguintes:

Criança 1



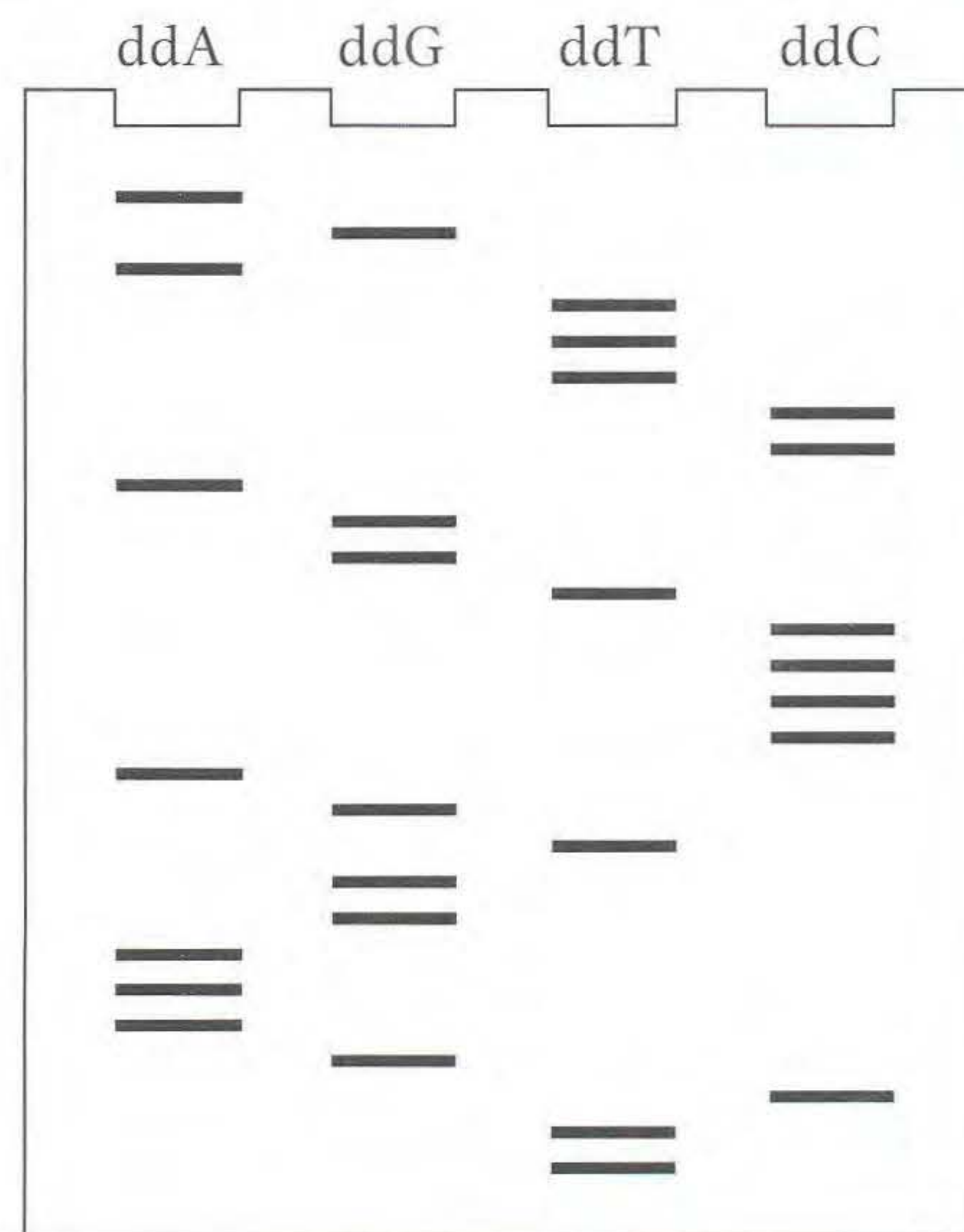


Para a criança 2, a atividade enzimática de cada estágio foi muito baixa e podia ser estimada apenas em cerca de 0,1 unidade nas idades 1, 2, 3 e 4.

- Para ambas as crianças, faça gráficos representando a expressão desenvolvimental do gene. (Marque totalmente ambos os eixos.)
- Como você pode explicar os níveis muito baixos de enzima ativa para a criança 2? (A degradação da proteína é apenas uma possibilidade.)
- Como você pode explicar a diferença na transferência de Southern para a criança 2 comparada com a da criança 1?
- Se apenas um gene mutante foi detectado nos estudos familiares de duas crianças, defina as crianças individuais como homozigotas ou heterozigotas para o gene *D*.

(O Problema 17 é cortesia de Joan McPherson. De A.J.F. Griffiths and J. McPherson, *1001 Principles of Genetics*. W.H. Freeman and Company, 1989.)

- Um fragmento clonado de DNA foi seqüenciado usando o método dideoxi. Uma parte do auto-radiograma do seqüenciamento do gel é representada aqui.



- Deduz a seqüência de nucleotídeos da cadeia de nucleotídeos do DNA sintetizada a partir do primer. Marque as pontas 5' e 3'.
 - Deduz a seqüência de nucleotídeos da cadeia de nucleotídeos do DNA usada como filamento-molde. Marque as pontas 5' e 3'.
 - Escreva a seqüência de nucleotídeos da dupla hélice de DNA (marque as pontas 5' e 3').
 - Quantas das seis matrizes de leitura estão "abertas" tanto quanto você consiga saber?
- O clone de cDNA para o gene humano codificando tirosinase foi marcado radioativamente e usado em uma análise de Southern do DNA genômico digerido com *EcoRI* de um camundongo tipo selvagem. Três fragmentos de camundongo foram encontrados sendo radioativos (estavam ligados à sonda). Quando foram usados camundongos albinos nessa análise de Southern, nenhum fragmento genômico ligou-se à sonda. Explique esses resultados em relação à natureza do tipo selvagem e alelos mutantes de camundongo.
 - Plantas transgênicas de tabaco foram obtidas, nas quais o plasmídeo vetor Ti foi destinado a inserir o gene de interesse mais um gene adjacente de resistência à canamicina. A herança da inserção cromossômica foi seguida testando a prole para resistência à canamicina. Duas plantas tipificam os resultados geralmente obtidos. Quando a planta 1 foi retrocruzada com tabaco tipo selvagem, 50% da prole eram resistentes à canamicina e 50% sensíveis. Quando a planta 2 foi retrocruzada com o tipo selvagem, 75% da prole eram resistentes à canamicina e 25% sensíveis. Qual deve ter sido a diferença entre as duas plantas transgênicas? O que você poderia prever sobre a situação quanto ao gene de interesse?
 - Uma mutação de fibrose cística em determinado heredograma é devida a um único par de nucleotídeos. Essa

mudança destrói um sítio de restrição *EcoRI* normalmente encontrado nessa posição. Como você usaria essa informação em uma consulta de membros dessa família sobre sua probabilidade de serem portadores? Cite os experimentos precisos necessários. Suponha que você descobre que uma mulher nessa família é portadora, e que ela é casada com um homem não-aparentado que também é heterozigoto para fibrose cística, mas, nesse caso, é uma mutação diferente no mesmo gene. Como você informaria esse casal sobre os riscos de um filho ter fibrose cística?

- 22.** A glicuronidase bacteriana converte uma substância incolor chamada X-Gluc em um pigmento azul-claro índigo. O gene para glicuronidase também funciona em plantas se elas receberem uma região promotora. Como você usaria esse gene como gene repórter para encontrar os tecidos nos quais um gene de planta que você clonou

é normalmente ativo? (Suponha que X-Gluc é facilmente captada pelos tecidos da planta.)

- 23.** A planta *Arabidopsis thaliana* foi transformada usando-se o plasmídeo Ti no qual havia sido inserido um gene de resistência à canamicina na região T-DNA. Duas colônias resistentes à canamicina (A e B) foram selecionadas, e foram geradas plantas a partir delas. As plantas se autopolinizaram, e os resultados foram os seguintes:

Planta A → $\frac{3}{4}$ da prole resistentes à canamicina
autofecundada $\frac{1}{4}$ da prole sensível à canamicina

Planta B → $\frac{15}{16}$ da prole resistentes à canamicina
autofecundada $\frac{1}{16}$ da prole sensível à canamicina

- a.** Desenhe os cromossomos relevantes da planta em ambas as plantas.
b. Explique as duas proporções diferentes.

Um Guia Resumido para Organismos-modelo

Escherichia coli • *Saccharomyces cerevisiae* • *Neurospora crassa* • *Arabidopsis thaliana* •
Caenorhabditis elegans • *Drosophila melanogaster* • *Mus musculus*

Este guia resumido reúne as principais características dos organismos-modelo relacionados à genética. Cada um dos sete organismos-modelo é mostrado em duas páginas; o formato é consistente, permitindo que os leitores comparem e contrastem as características dos organismos-modelo. Cada tratamento enfoca as características do organismo que o tornaram útil como modelo; as técnicas especiais que foram desenvolvidas para estudar o organismo; e as principais contribuições que os estudos do organismo deram para nossa compreensão da genética. Embora muitas diferenças sejam aparentes, os enfoques gerais da análise genética são similares, mas têm que ser ajustados para considerar o ciclo de vida individual, nível de ploidia, tamanho e forma, e propriedades genômicas, tais como a presença de plasmídeos naturais e transposons.

Os organismos-modelo sempre estiveram na fronteira da Genética. Inicialmente, no desenvolvimento histórico de um organismo-modelo, um pesquisador seleciona o organismo devido a uma característica que se ajusta particularmente bem ao estudo de um processo genético no qual o pesquisador está interessado. O conselho dos últimos cem anos foi: “Escolha bem seu organismo.” Por exemplo, o fungo ascomi-ceto, tal como *Saccharomyces cerevisiae* e *Neurospora crassa*, é bem adequado ao estudo de processos meióticos, tal como o crossing over, devido à sua característica única, o asco, de manter juntos os produtos de uma única meiose.

Espécies diferentes tendem a apresentar processos mar-cantemente similares, mesmo através de grandes grupos, tais como os eucariontes. Assim, podemos razoavelmente esperar que o que se aprende em uma espécie pode ser pelo menos parcialmente aplicado a outras. Em particular, os geneticistas mantiveram os olhos abertos para novos achados de pesquisa que podem ser aplicados a nossa própria espécie. Comparado com outras espécies, os humanos são relativamente difíceis de estudar em âmbito genético, e, assim, os avanços em genética humana devem-se, em grande parte, a mais de um século de trabalho em organismos-modelo.

Todos os organismos-modelo têm bem mais que uma característica útil para o estudo genético ou outros biológicos. Assim, após um organismo-modelo ter sido desenvolvido por algumas pessoas com interesses específicos, ele então atua como um núcleo para o desenvolvimento de uma comunidade de pesquisa, um grupo de pesquisadores com interesse em várias características de um organismo-modelo em particular. Existem comunidades de pesquisa organizadas para todos os organismos-modelo mencionados neste resumo. As pessoas nessas comunidades estão em contato umas com as outras regularmente, compartilham suas linhagens mutantes e, em geral, encontram-se pelo menos anualmente em conferências que podem atrair milhares de pessoas. Tais comunidades possibilitam o fornecimento de serviços importantes, tais como bancos de dados de informações sobre pesquisas, técnicas, estoques de genética, clones, bibliotecas de DNA e seqüências genômicas.

Outra vantagem para um pesquisador individual em pertencer a tal comunidade é que ele desenvolve “um *feeling* para o organismo” (uma frase da geneticista de milho e ganhadora de um Nobel, Barbara McClintock). Essa idéia é difícil de transmitir, mas implica uma compreensão dos caminhos gerais de um organismo. Nenhum processo vivo ocorre no isolamento, e, assim, conhecer os caminhos gerais de um organismo é geralmente benéfico para compreender um processo e interpretá-lo em seu contexto apropriado.

À medida que aumenta o banco de dados para cada organismo-modelo (o que está atualmente sendo feito em grande velocidade graças à genômica), os geneticistas são mais e mais capazes de ter uma visão holística, englobando os trabalhos integrados de todas as partes da constituição do organismo. Desse modo, os organismos-modelo tornam-se não só modelos para processos isolados, mas também modelos de processos integrados de vida. O termo *biologia de sistemas* é usado para descrever esse enfoque holístico.



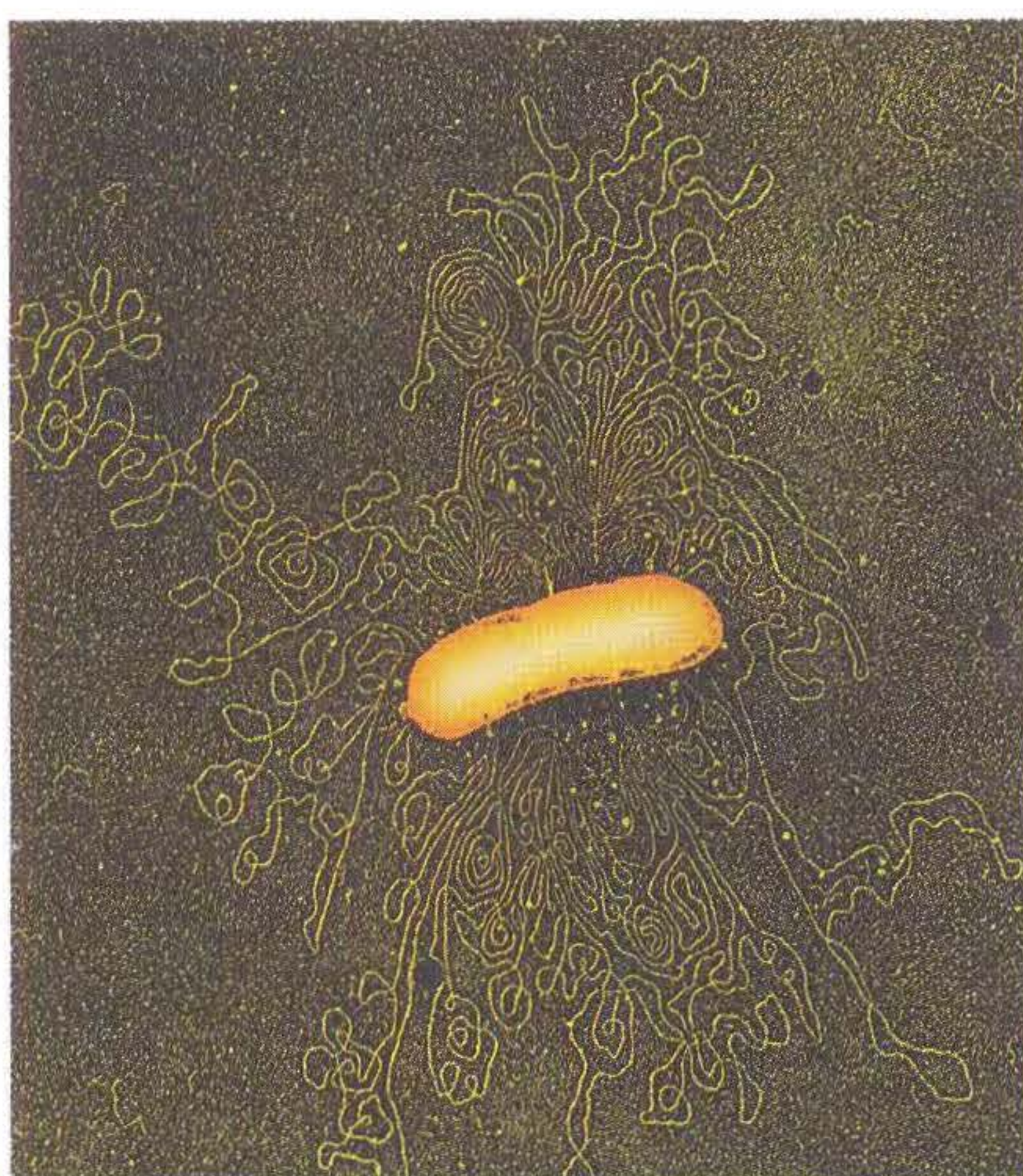
Escherichia coli

Organismo-chave para estudar:

- Transcrição, tradução, replicação, recombinação
- Mutação
- Regulação gênica
- Tecnologia de DNA recombinante

"Estatísticas Vitais" Genéticas

Tamanho do genoma:	4,6 Mb
Cromossomos:	1, circular
Número de genes:	4.000
Porcentagem com homólogos humanos:	8%
Tamanho médio dos genes:	1 kb, sem íntrons
Transposons:	Específicos de linhagem, ~60 cópias por genoma
Genoma seqüenciado em:	1997

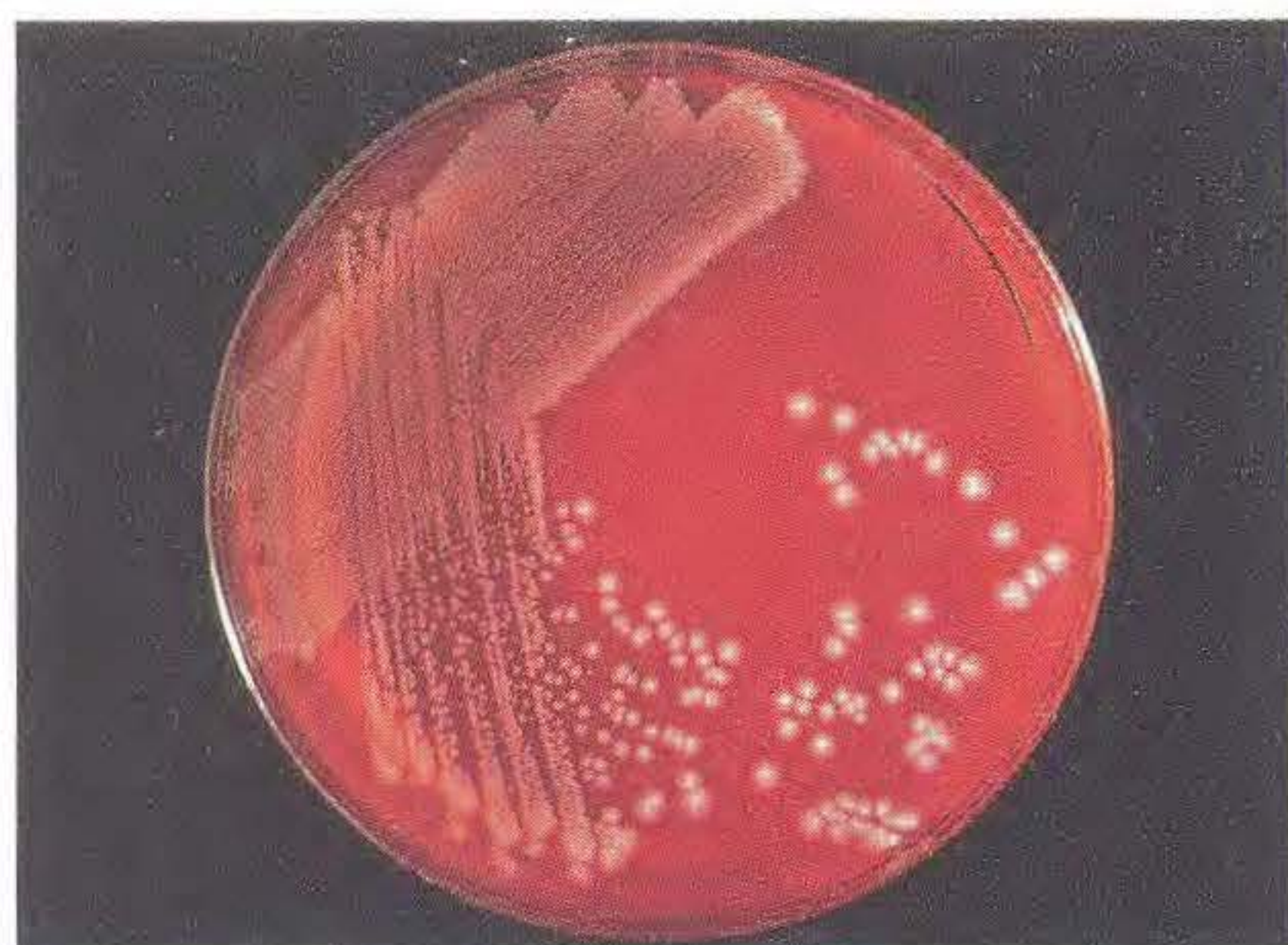


A bactéria unicelular *Escherichia coli* é amplamente conhecida como um patógeno causador de doenças, uma fonte de envenenamento alimentar e doença intestinal. Entretanto, essa reputação negativa é indevida. Embora algumas linhagens de *E. coli* sejam prejudiciais, outras são naturais e residentes essenciais do intestino humano. Como organismo-modelo, as linhagens de *E. coli* têm um papel indispensável na análise genética. Nos anos de 1940, vários grupos começaram a investigar a genética de *E. coli*. A necessidade era de um organismo simples que pudesse ser cultivado de modo barato para produzir um grande número de bactérias, permitindo encontrar e analisar eventos genéticos raros. Como a *E. coli* pode ser obtida do intestino humano e é pequena e fácil de cultivar, ela foi uma escolha natural. O trabalho em *E. coli* definiu o começo do esclarecimento da "caixa preta" em Genética: pela seleção e análise dos mutantes, os trabalhos de processos celulares podiam ser deduzidos, muito embora uma célula individual fosse muito pequena para ser vista.

Genoma de *E. coli*. Micrografia eletrônica do genoma da bactéria *E. coli*, liberado da célula por choque osmótico. [Dr. Gopal Murti/Science Photo Library/Photo Researchers.]

Características especiais

Grande parte do sucesso de *E. coli* como um organismo-modelo pode ser atribuído a duas estatísticas: seu tamanho celular de 1 μm e o tempo de geração de 20 minutos. (A replicação do cromossomo leva 40 minutos, mas muitas forquilhas de replicação permitem que a célula se divida em 20 minutos.) Conseqüentemente, esse procarionte pode ser cultivado em grandes números, uma característica que permite aos geneticistas identificar mutações e outros eventos genéticos raros, tais como recombinantes intragênicos. *E. coli* também é muito fácil de cultivar. Quando as células são espalhadas em placas de meio nutritivo, cada célula divide-se *in situ* e forma uma colônia viável. Alternativamente, banhos de células podem ser cultivadas em cultura líquida. Fenótipos tais como tamanho da colônia, resistência a drogas, habilidade de obter energia de determinada fonte de carbono e produção de corantes tomam o lugar dos fenótipos morfológicos da genética de eucariontes.



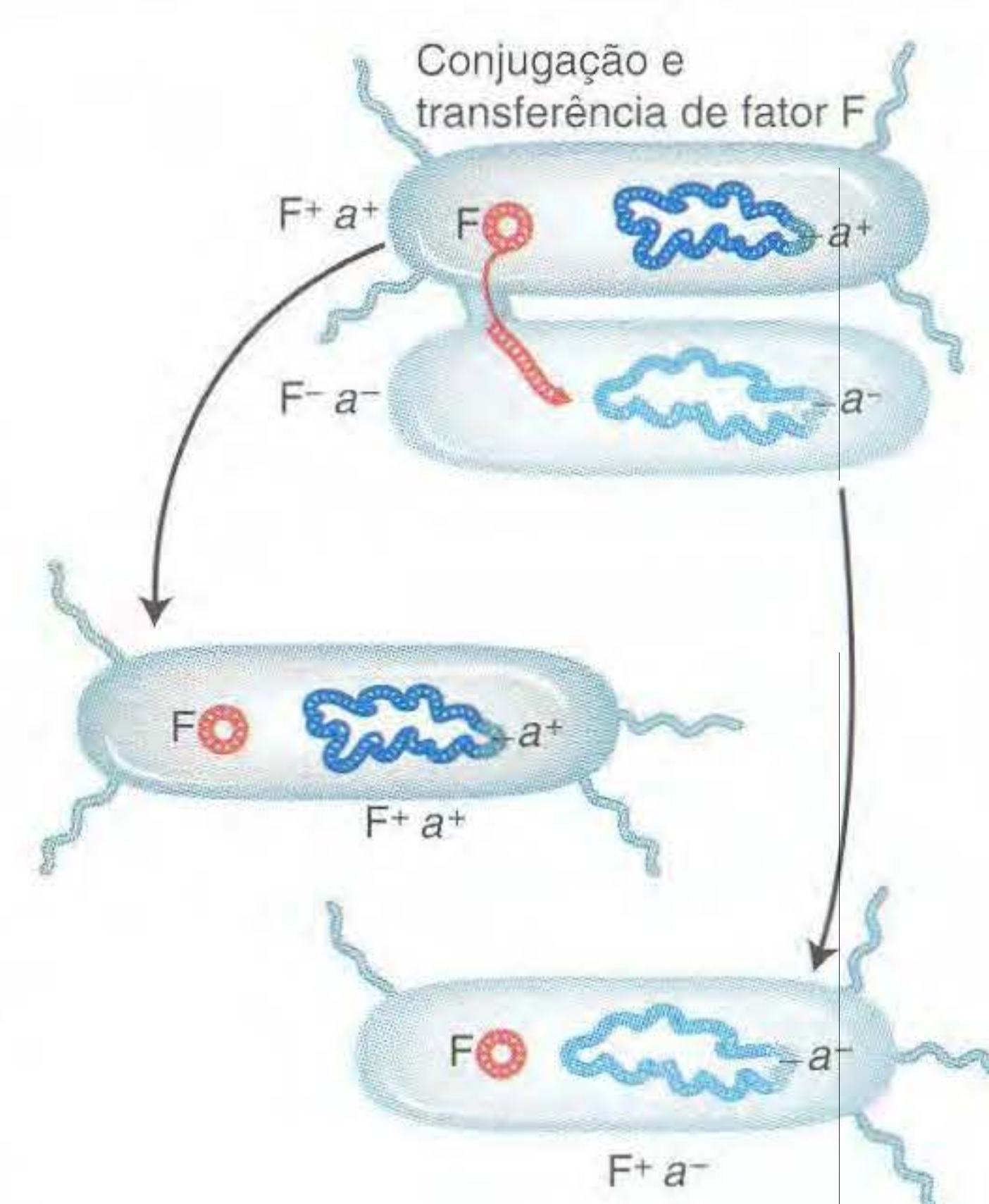
Colônias bacterianas. [Biophoto Associates/Science Source/Photo Researchers.]

Ciclo de Vida

A *Escherichia coli* reproduz-se assexualmente por divisão simples; seu genoma haplóide replica-se e é repartido na divisão celular. Nos anos de 1940, Joshua Lederberg e Edward Tatum descobriram que *E. coli* também tem um tipo de ciclo sexual no qual as células de "sexos" geneticamente diferenciados se fundem e trocam parcial ou totalmente os seus genomas, às vezes levando à recombinação (veja Capítulo 5).

"Machos" podem converter "fêmeas" em machos pela transmissão de um plasmídeo particular. Esse plasmídeo de DNA extragenômico circular com 100 kb, chamado F, determina um tipo de "masculinidade". As células F^+ agindo como "doadores" masculinos transmitem uma cópia do plasmídeo F para uma célula receptora. O plasmídeo F pode integrar-se ao cromossomo para formar um tipo de célula Hfr, que transmite o cromossomo linearmente para receptoras F^- . Outros plasmídeos são encontrados em *E. coli* na natureza. Alguns levam genes cujas funções equipam a célula para a vida em ambientes específicos. Os plasmídeos R que levam genes de resistência a drogas são exemplos.

Duração do ciclo de vida: 20 minutos



Os geneticistas também tiraram proveito de alguns elementos genéticos únicos associados a *E. coli*. Os plasmídeos bacterianos e os fagos são usados como vetores para clonar os genes de outros organismos dentro de *E. coli*. Os elementos de transposição de *E. coli* são equipados para perturbar genes em DNA eucariótico clonado. Tais elementos bacterianos são participantes-chave na tecnologia do DNA recombinante.

Análise genética

Mutantes espontâneos de *E. coli* mostram uma variedade de mudanças de DNA, variando desde simples substituições de bases até inserções de elementos de transposição. O estudo de raras mutações espontâneas em *E. coli* é factível porque grandes populações podem ser triadas. Entretanto, também podem ser usados mutágenos para aumentar as frequências de mutação.

Para obter fenótipos mutantes específicos que podem representar defeitos em um processo em estudo, as triagens ou seleções devem ser criadas. Por exemplo, mutações nutricionais e mutações que conferem resistência a drogas ou fagos podem ser obtidas em placas suplementadas com substâncias específicas, drogas ou fagos. As mutações nulas de qualquer gene essencial resultarão em falta de crescimento. Essas mutações podem ser selecionadas adicionando-se penicilina (uma droga antibacteriana isolada de um fungo), que mata as células em divisão, mas não os mutantes que não crescem. Para mutações letais condicionais, pode ser usado o plaqueamento réplica: colônias mutadas na placa master são transferidas por uma almofada de feltro para outras placas, que são então sujeitas a algum ambiente tóxico. As mutações que afetam a expressão de um gene específico de interesse podem ser triadas ligando-o a um gene repórter tal como o gene *lacZ*, cujo produto proteico pode fazer um corante azul, ou o gene *GFP*, cujo produto fluoresce quando exposto à luz de determinado comprimento de onda.

Após ter sido obtido um conjunto de mutantes que afetam o processo de interesse, as mutações são distribuídas pelos seus genes por recombinação e complementação. Esses genes são clonados e seqüenciados para obter indícios da função. A mutagênese-alvo pode ser usada para ajustar as mudanças mutacionais a posições específicas da proteína (veja Capítulo 13).

Em *E. coli*, os cruzamentos são usados para mapear mutações e produzir genótipos celulares específicos (veja Capítulo 5). Os recombinantes são feitos misturando células Hfr (tendo um plasmídeo F integrado) e células F⁻. Em geral, um doador Hfr transmite parte do genoma bacteriano, formando um merozigoto temporário no qual ocorre a recombinação. Os cruzamentos de Hfr podem ser usados para fazer o mapeamento pelo tempo de entrada do marcador ou pela frequência de recombinantes. Pela transferência de derivados F' levando genes doadores para F⁻, é possível fazer diplóides parciais estáveis para estudar a interação gênica ou dominância.

Técnicas de Modificação Genética

Mutagênese padrão:

Substâncias químicas e radiação	Mutações somáticas aleatórias
Transposons	Inserções somáticas aleatórias

Transgênese:

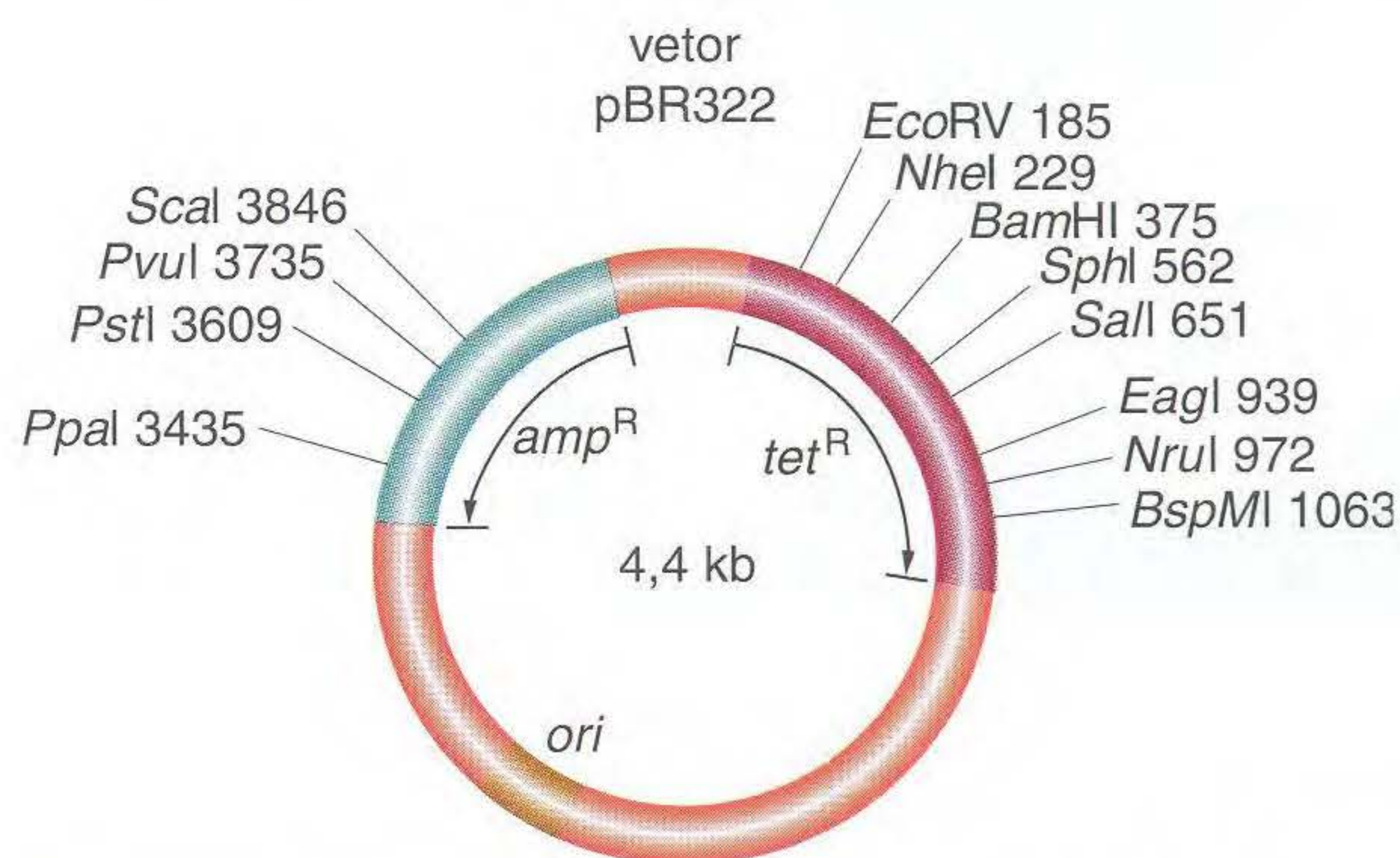
Em plasmídeo vetor	Livre ou integrado
Em fago vetor	Livre ou integrado
Transformação	Integrada

Nocautes gênicos direcionados:

Alelo nulo em vetor	Substituição de gene por recombinação
Alelo construído em vetor	Mutagênese sítio-dirigida por substituição gênica

Engenharia genética

Transgênese. *E. coli* tem um papel importante em introduzir transgenes em outros organismos (veja Capítulo 20). É o organismo padrão usado para clonagem de genes de qualquer organismo. Os plasmídeos de *E. coli* ou bacteriófagos são usados como vetores, levando a seqüência de DNA a ser clonada. Esses vetores são introduzidos em uma bactéria por transformação, se um plasmídeo, ou por transdução, se um fago, que se replica no citoplasma. Os vetores são especialmente modificados para incluir sítios de clonagem únicos que podem ser cortados por uma variedade de enzimas de restrição. Outros vetores de “transporte” são destinados a mover fragmentos de DNA de levedura (“*E. coli* eucariótica”) para *E. coli* pela sua maior facilidade de manipulação genética, e então de volta à levedura para avaliação fenotípica.



Um plasmídeo construído como vetor de clonagem de DNA. A inserção bem-sucedida de um gene exógeno no plasmídeo é detectada pela inativação de genes de resistência a drogas (*tet^R* ou *amp^R*). São identificados os sítios de restrição.

Nocautes gênicos direcionados. Um conjunto completo de nocautes gênicos está sendo acumulado. Em um procedimento, um transposon de resistência à canamicina é introduzido em um gene clonado *in vitro* (usando uma transposase). A construção é transformada, e as colônias resistentes são nocautes produzidos por recombinação homóloga.

Contribuições principais

Estudos pioneiros de genética como um todo foram feitos em *E. coli*. Talvez o maior triunfo foi a elucidação do código genético universal de 64 códons, mas essa conquista está longe de ser a única na lista de conquistas atribuíveis a esse organismo. Outros fundamentos de genética que foram primeiro demonstrados em *E. coli* incluem a natureza espontânea da mutação (o teste de flutuação, Capítulo 15), os vários tipos de mudanças de bases que causam as mutações, e a replicação semiconservativa do DNA (o experimento de Meselson e Stahl, Capítulo 7). Esta bactéria ajudou a abrir novas áreas da genética, tais como a regulação gênica (o óperon *lac*, Capítulo 10) e a transposição de DNA (os elementos IS, Capítulo 14). Finalmente, a tecnologia do DNA recombinante foi inventada em *E. coli*, e o organismo ainda tem um papel central nessa tecnologia hoje em dia.

Outras áreas de contribuição

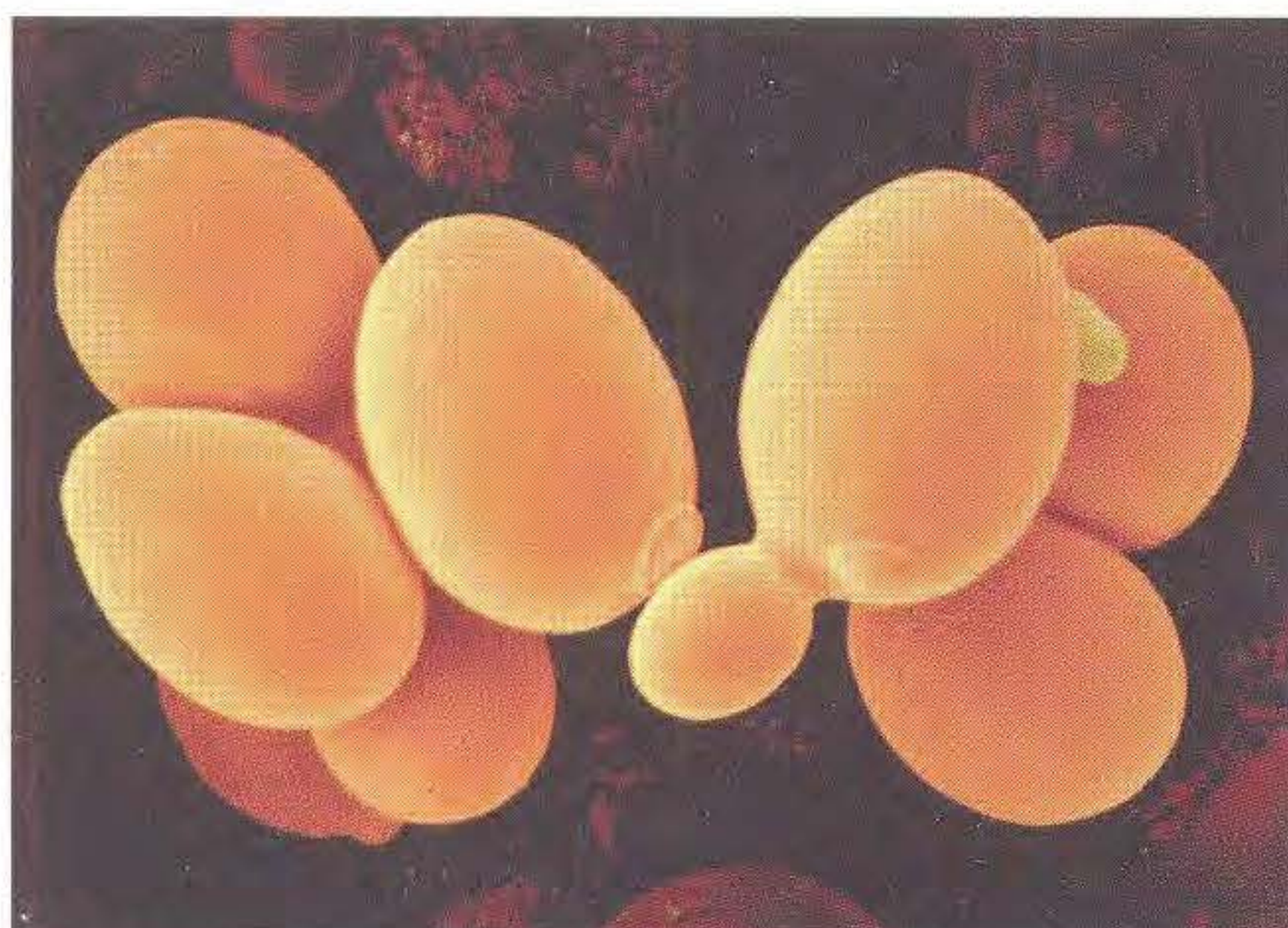
- Metabolismo celular
- Supressores não-sensíveis
- Colinearidade do gene e polipeptídeo
- O óperon
- Resistência a drogas baseada em plasmídeo
- Transporte ativo



Saccharomyces cerevisiae

Organismo-chave para estudar:

- Genômica
- Biologia de sistema
- Controle genético do ciclo celular
- Transdução de sinal
- Recombinação
- Tipo reprodutivo
- Herança mitocondrial
- Interação gênica; diíbrido



O ascomiceto *S. cerevisiae*, a “levedura do pão”, ou simplesmente “levedura”, tem sido a base das indústrias de pão e cerveja desde a antiguidade. Na natureza, ela provavelmente cresce nas superfícies das plantas, usando exsudados como nutrientes, embora seu nicho preciso ainda seja um mistério. Embora as linhagens laboratoriais sejam principalmente haplóides, as células na natureza podem ser diplóides ou poliplóides. Em aproximadamente 70 anos de pesquisas genéticas, as leveduras tornaram-se “as *E. coli* dos eucariontes”. Como é haplóide, unicelular e forma colônias compactas nas placas, ela pode ser tratada do mesmo modo que uma bactéria. Entretanto, tem meiose eucariótica, ciclo celular e mitocôndrias, e essas características estiveram no centro da história de sucesso de leveduras.

Células de levedura, *Saccharomyces cerevisiae*.

Características especiais

Como um organismo-modelo, as leveduras combinam o melhor dos dois mundos: têm muito da conveniência de uma bactéria, mas com as características-chave de um eucarionte. As células de leveduras são pequenas (10 μm) e completam seu ciclo celular apenas em 90 minutos, permitindo que sejam produzidas em grandes números em um tempo curto. Como as bactérias, as leveduras podem ser cultivadas em grandes banhos de meio líquido que são continuamente agitados. Como qualquer bactéria, as leveduras produzem colônias quando plaqueadas em meio de ágar, podem ser triadas quanto a mutações e ter placa réplica. De modo tipicamente eucariótico, as leveduras têm um ciclo de divisão mitótico, sofrem meiose e contêm mitocôndrias abrigando um único genoma pequeno. As células de leveduras podem respirar anaerobicamente usando o ciclo de fermentação, e, assim, podem fazê-lo sem mitocôndrias, permitindo que os mutantes mitocondriais sejam viáveis.

Análise genética

Fazer cruzamentos em leveduras é bem simples. As linhagens de tipo reprodutivo oposto são simplesmente misturadas em um meio apropriado. Os diplóides resultantes a/α são induzidos a sofrer meiose usando um meio especial de esporulação. Os pesquisadores podem isolar ascósporos de uma única tétrade usando uma máquina chamada micromanipulador. Eles também têm a opção de sintetizar a/a ou diplóides α/α para fins especiais ou criar diplóides parciais usando plasmídeos especialmente construídos.

Devido a uma grande gama de mutantes de leveduras e construções de DNA estarem disponíveis na comunidade de pesquisa, podem ser feitas linhagens para fins especiais para triagem e seleção cruzando vários tipos de leveduras. Adicionalmente, novos alelos mutantes podem ser mapeados cruzando com linhagens contendo uma gama de marcadores de DNA ou fenotípicos da posição de mapa conhecida.

A disponibilidade de células, tanto haplóides quanto diplóides, fornece flexibilidade a estudos mutacionais. As células haplóides são convenientes para seleções em larga escala ou triagens porque os fenótipos mutantes são expressos diretamente. As células diplóides

“Estatísticas Vitais” Genéticas

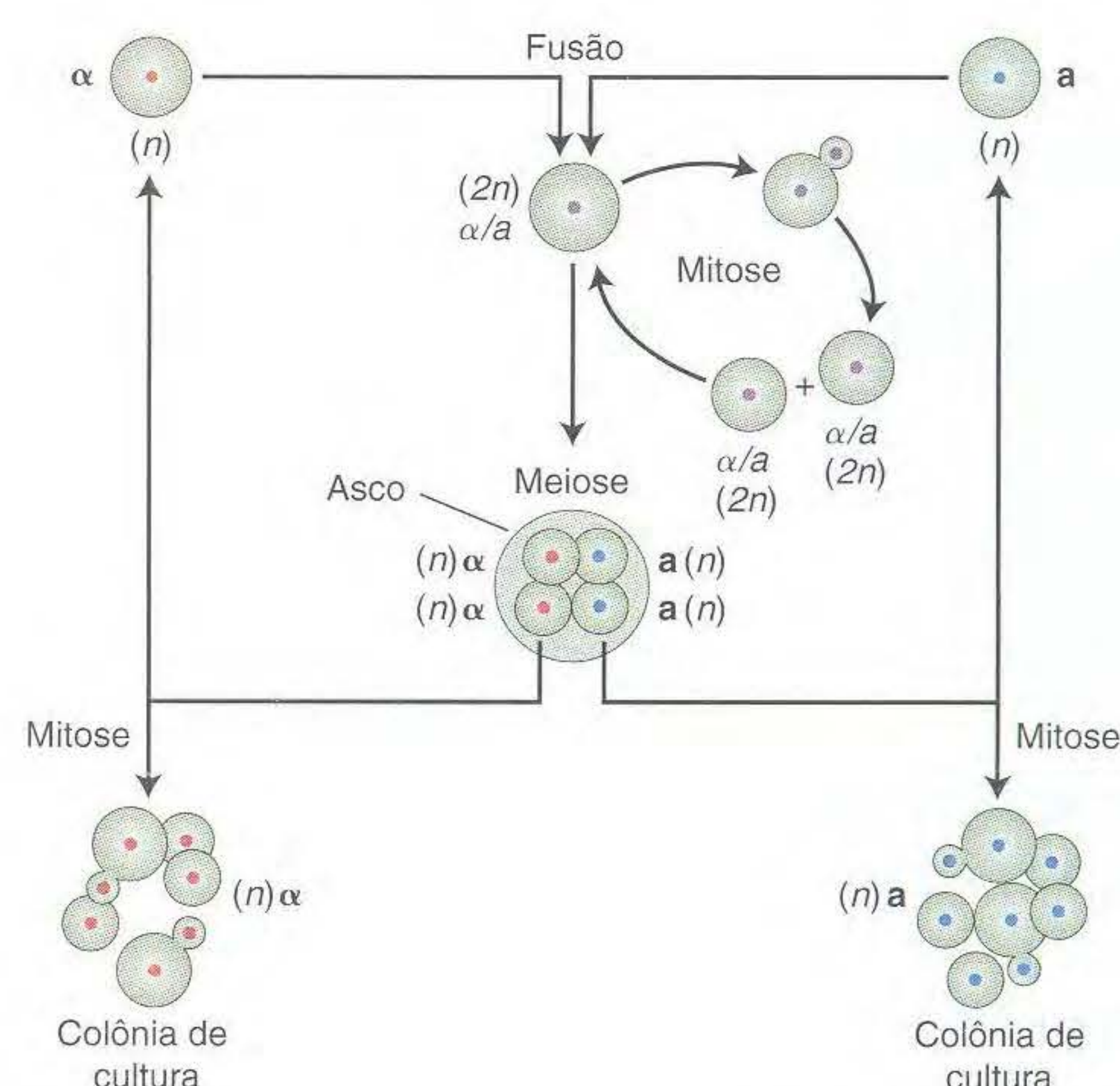
Tamanho do genoma:	12 Mb
Cromossomos:	$n = 16$
Número de genes:	6.000
Porcentagem com homólogos humanos:	25%
Tamanho médio dos genes:	1,5 kb, 0,03 íntron/gene
Transposons:	Pequena proporção de DNA
Genoma seqüenciado em:	1996

são convenientes para obter mutações dominantes, abrigar mutações letais, fazer testes de complementação e explorar interação gênica.

Ciclo de Vida

As leveduras são uma espécie unicelular com um ciclo de vida simples, que consiste nas fases sexual e assexual. A fase assexual pode ser haplóide ou diplóide. Uma célula divide-se assexualmente por brotamento: uma célula-mãe elimina um broto para o qual é passado um dos núcleos que resultam da mitose. Para a reprodução sexual, existem dois tipos principais, determinando os alelos $MAT\alpha$ e $MATa$. Quando as células diplóides de tipos reprodutivos diferentes se unem, elas formam uma célula diplóide, que pode dividir-se mitoticamente e sofrer divisão meiótica. Os produtos da meiose são uma tétrade não-linear de quatro ascósporos.

Duração total do ciclo de vida: 90 minutos para completar o ciclo celular



Técnicas de Manipulação Genética

Mutagênese padrão:

Substâncias químicas e radiação	Mutações somáticas aleatórias
Transposons	Inserções somáticas aleatórias

Transgênese:

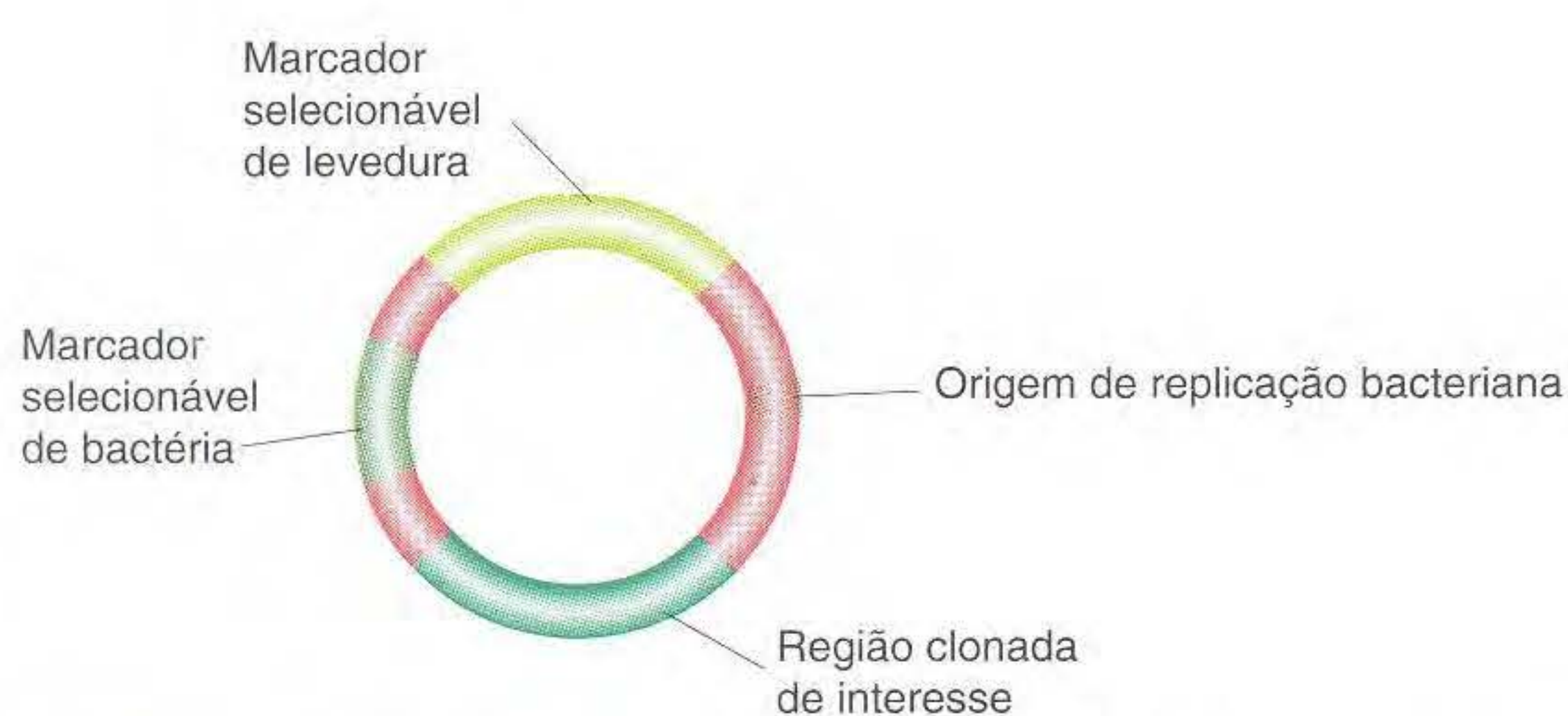
Plasmídeo integrativo	Inserção por recombinação homóloga
Plasmídeo replicativo	Pode se replicar autonomamente (2 μ ou origem ARS de replicação)
Cromossomo artificial de levedura	Replica-se e segrega-se como um cromossomo
Vetor de transporte	Pode se replicar em leveduras ou <i>E. coli</i>

Nocautes gênicos direcionados:

Substituição gênica	Recombinação homóloga substitui alelo tipo selvagem por cópia nula
---------------------	--

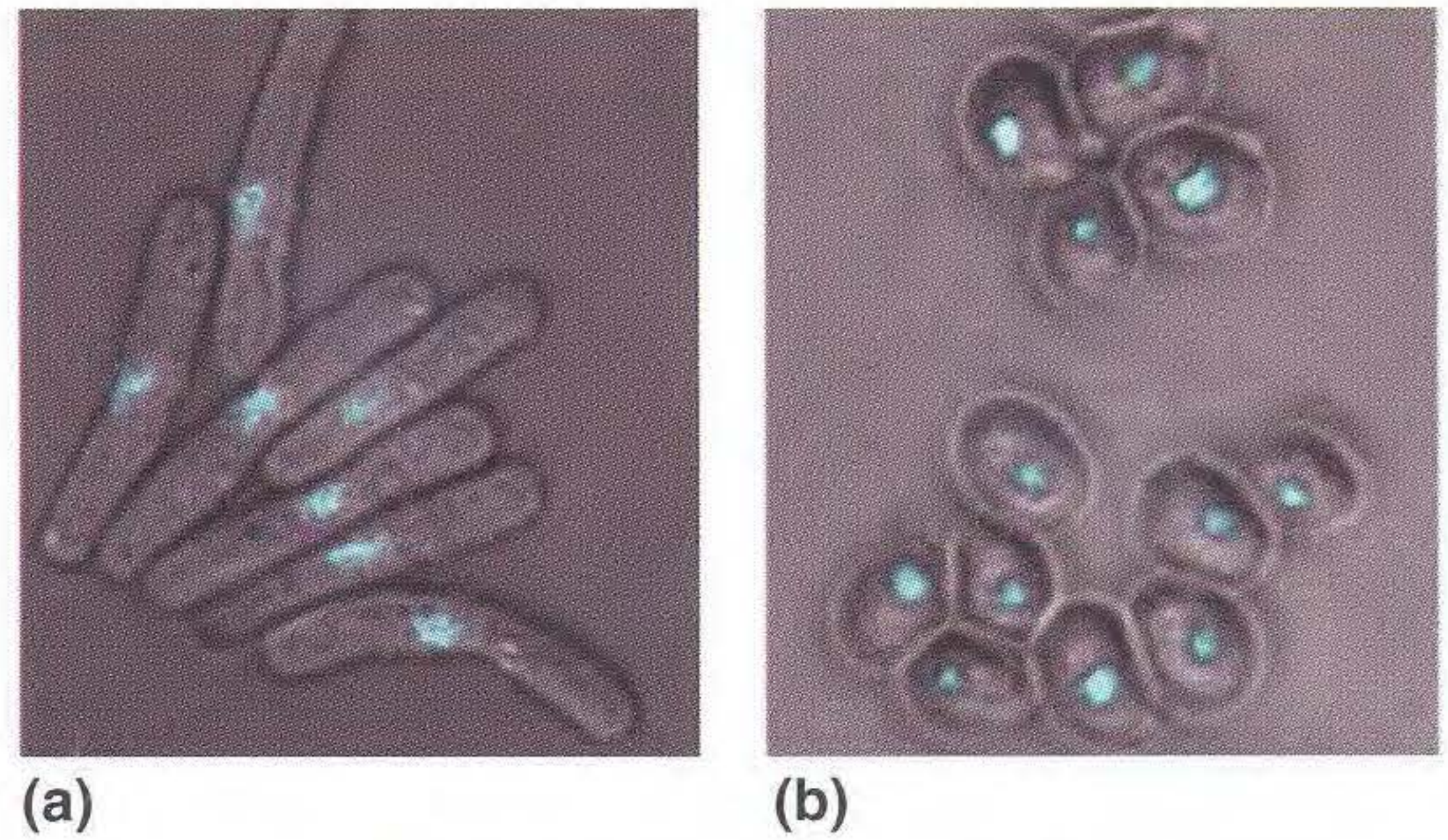
Engenharia genética

Transgênese. A levedura de brotamento fornece mais oportunidades para manipulação genética do que qualquer outro eucarionte (veja Capítulo 20). O DNA exógeno é captado facilmente pelas células cujas paredes celulares foram parcialmente removidas por digestão enzimática ou abrasão. Vários tipos de vetores estão disponíveis (veja Técnicas de Manipulação Genética). Para que um plasmídeo possa se replicar livre de cromossomos, ele deve conter uma origem normal de replicação de levedura (ARS) ou uma origem de replicação de um plasmídeo de 2 μ m encontrado em alguns isolados de levedura. O vetor mais elaborado, o cromossomo artificial de levedura (YAC), consiste em uma ARS e dois telômeros. Um YAC pode levar grandes insertos transgênicos, que são então herdados do mesmo modo que os cromossomos mendelianos. Os YAC têm sido vetores importantes na clonagem e seqüenciamento de grandes genomas, tais como o genoma humano.



Um vetor simples de levedura. Esse tipo de vetor é chamado de plasmídeo integrativo de levedura (Ylp).

Nocautes direcionados. A mutagênese de transposon (marcação de transposon) pode ser feita introduzindo DNA de levedura em *E. coli* em um vetor de transporte; os transposons bacterianos integram-se ao DNA de levedura, inativando o funcionamento do gene. O vetor de transporte é então transferido de volta à levedura, e os mutantes marcados substituem cópias tipo selvagem por recombinação homóloga. Os nocautes gênicos também podem ser obtidos substituindo os alelos tipo selvagem por uma cópia nula construída por recombinação homóloga. Usando essas técnicas, os pesquisadores construíram sistematicamente um conjunto completo de linhagens nocauteadas de leveduras (cada uma levando um nocaute diferente) para cada gene a nível fenotípico.



Mutantes de ciclo celular. (a) Mutantes que se alongam sem se dividir. (b) Mutantes que param sem brotamento.

Contribuições principais

Graças a uma combinação de uma boa genética e uma boa bioquímica, os estudos de leveduras fizeram contribuições substanciais a nossa compreensão do controle genético dos processos celulares. **Ciclo celular.** A identificação dos genes de divisão celular por seus mutantes sensíveis à temperatura (mutantes *cdc*) levou a um modelo poderoso para o controle genético da divisão celular. Os fenótipos Cdc diferentes revelam os componentes da maquinaria necessária para executar etapas específicas na progressão do ciclo celular. Esse trabalho foi útil para compreender os controles anormais da divisão celular que levaram ao câncer humano.

Recombinação. Muitas das importantes idéias dos atuais modelos moleculares do crossing over (tais como o modelo da quebra bifilamentar) são baseadas na análise de tétrades da conversão gênica em leveduras (veja Capítulo 15). A conversão gênica (proporções alélicas aberrantes tais como 3:1) é muito comum em genes de leveduras, dando um apropriado grande número de dados para quantificar as características importantes desse processo.

Interações gênicas. As leveduras abriram caminho no estudo das interações gênicas. As técnicas de genética tradicional foram usadas para revelar padrões de epistasia e supressão, que sugerem interações gênicas (veja Capítulo 6). O sistema de plasmídeo díbrido para encontrar interações de proteínas foi desenvolvido em leveduras e gerou mapas complexos de interação que representam o início da biologia de sistemas (veja Capítulo 13). Mutantes duplos letal-letal sintéticos, criados por entrecruzamento de dois mutantes únicos viáveis, também são usados para plotar redes de interação (veja Capítulo 6).

Genética mitocondrial. Mutantes com mitocôndrias defeituosas são reconhecíveis como colônias muito pequenas chamadas de “petites”. A disponibilidade desses petites e de outros mutantes mitocondriais possibilitou a primeira análise detalhada da estrutura do genoma mitocondrial e o funcionamento em qualquer organismo.

Genética do tipo reprodutivo. Os alelos *MAT* de levedura foram os primeiros genes de tipos reprodutivos a serem caracterizados a nível molecular. Curiosamente, as leveduras sofrem mudanças espontâneas de um tipo reprodutivo para outro. Uma cópia “reserva” silenciosa do alelo *MAT* oposto, residente em outro local no genoma, entra no locus do tipo reprodutivo, substituindo o alelo residente por recombinação homóloga. As leveduras forneceram um dos modelos centrais para a transdução do sinal durante a detecção e resposta aos hormônios reprodutivos do tipo oposto.

Outras áreas de contribuição

- A genética da mudança entre o crescimento tipo levedura e o filamentososo
- Genética da senescência



Neurospora crassa

Organismo-chave para estudar:

- Genética do metabolismo e captação
- Genética do crossing over e meiose
- Citogenética de fungos
- Crescimento polar
- Ritmo circadiano
- Interações entre núcleo e mitocôndrias

"Estatísticas Vitais" Genéticas

Tamanho do genoma:	43 Mb
Cromossomos:	7 autossomos ($n = 7$)
Número de genes:	10.000
Percentagem com homólogos humanos:	6%
Tamanho médio dos genes:	1,7 kb, 1,7 íntrons/gene
Transposons:	raros
Genoma seqüenciado em:	2003



Neurospora crassa crescendo em cana-de-açúcar.

Neurospora crassa, o fungo laranja do pão, foi um dos primeiros micróbios eucarióticos a serem adotados pelos geneticistas como organismo-modelo. Como as leveduras, esse fungo foi originalmente escolhido devido à sua haploidia, seu ciclo de vida simples e rápido e à facilidade com a qual pode ser cultivado. De significado particular é o fato de que crescerá em um meio com um conjunto definido de nutrientes, possibilitando estudar o controle genético da química celular. Na natureza, é encontrado em muitas partes do mundo crescendo em vegetais mortos. Como o fogo ativa seus ascospores dormentes, é facilmente coletado após queimadas — por exemplo, sob árvores queimadas e em campos de cultivos tais como de cana, que são rotineiramente queimados antes da colheita.

Características especiais

A *Neurospora* tem o recorde de velocidade para os fungos porque cada hifa cresce mais de 10 cm por dia. Esse crescimento rápido, combinado com seu ciclo de vida haplóide e habilidade de crescer em meio definido, tornou-a o organismo de escolha para o estudo de genética bioquímica da nutrição e captação de nutrientes.

Outra característica única da *Neurospora* (e fungos correlatos) permite aos geneticistas traçarem as etapas de meioses únicas. Os quatro produtos haplóides de uma meiose ficam juntos em um saco chamado asco. Cada um dos quatro produtos da meiose sofre outra divisão mitótica, resultando em uma óctade linear de oito ascospores (veja Capítulo 3). Essa característica torna a *Neurospora* um sistema ideal no qual estudar o crossing over, conversão gênica, rearranjos cromossômicos, não-disjunção meiótica e o controle genético da própria meiose. Os cromossomos, embora pequenos, são facilmente visíveis, e, assim, os processos meióticos podem ser estudados tanto a nível genético quanto cromossômico. Assim, em *Neurospora*, foram feitos estudos fundamentais sobre os mecanismos subjacentes a esses processos (veja Capítulo 4).

Análise genética

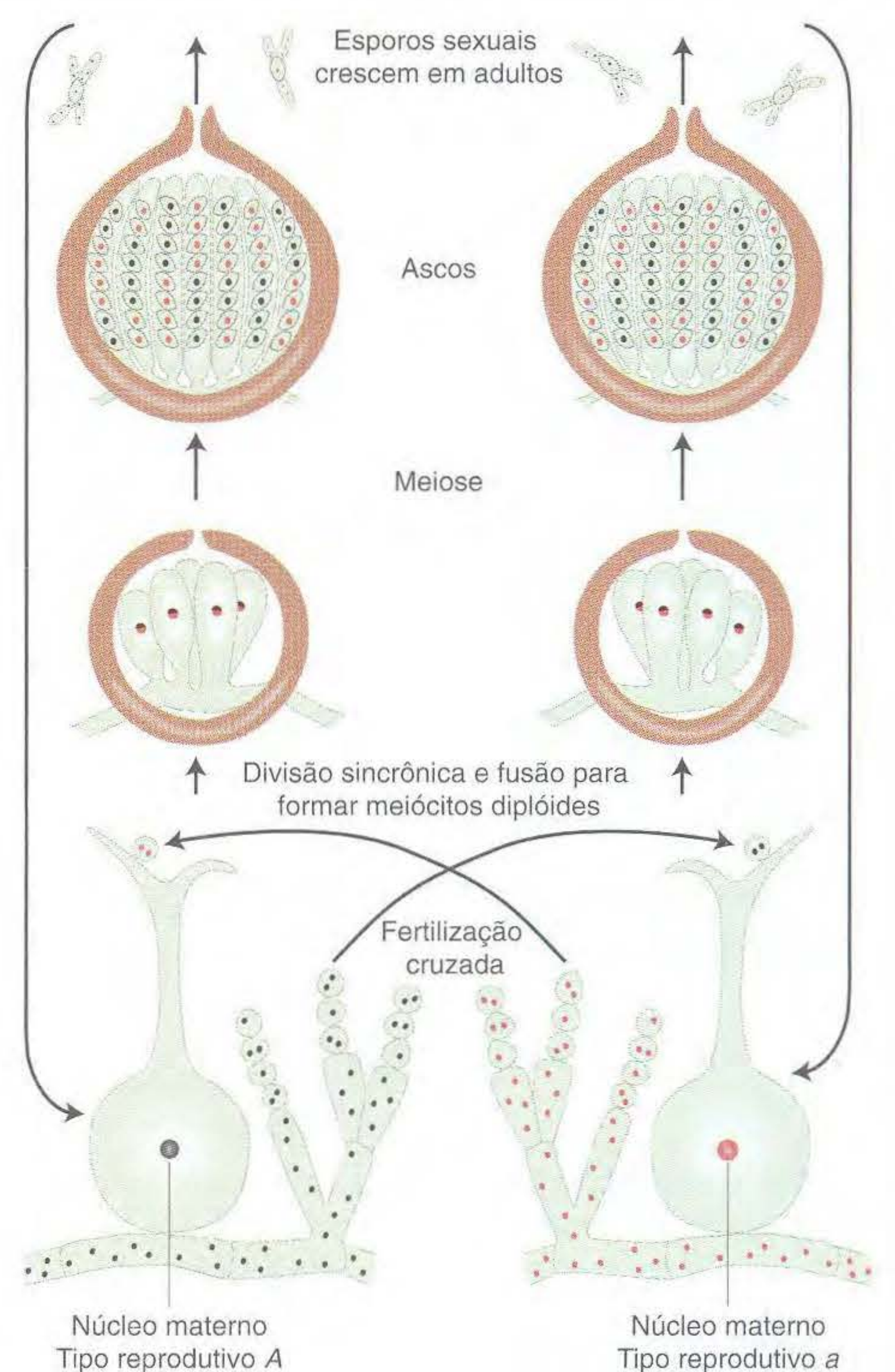
A análise genética é direta (veja Capítulo 3). Os centros de estoque fornecem uma ampla gama de mutantes que afetam todos os aspectos da biologia do fungo. Os genes de *Neurospora* podem ser facilmente mapeados se cruzados com um banco de linhagens com loci mutantes conhecidos ou alelos de RFLP conhecidos. As linhagens de tipo reprodutivo oposto são cruzadas simplesmente se cultivadas juntas. Um geneticista com uma agulha pode pegar

Ciclo de Vida

N. crassa tem um ciclo de vida eucariótico haplóide. Um esporo sexual haplóide (chamado de conídio) germina para produzir um tubo germinativo que se estende até a ponta. Progressivamente, o crescimento da ponta e a ramificação produzem uma massa de filamentos ramificados (chamados de hifas), que formam uma colônia compacta no meio de crescimento. Como as hifas não têm paredes cruzadas, uma colônia é essencialmente uma célula contendo muitos núcleos haplóides. A colônia brota de milhões de esporos assexuais, que podem dispersar-se no ar e repetir o ciclo assexual.

No ciclo sexual de *N. crassa*, existem dois tipos reprodutivos de aspecto idêntico, MAT-A e MAT-a, que podem ser vistos simplesmente como "sexos". Como em leveduras, os dois tipos reprodutivos são determinados por dois alelos de um gene. Quando as colônias de tipos reprodutivos diferentes entram em contato, suas paredes celulares e núcleos se fundem. Surgem muitos núcleos diplóides transitórios, cada um dos quais sofre meiose, produzindo uma óctade de ascospores. Os ascospores germinam e produzem colônias exatamente como as produzidas por esporos assexuais.

Duração do ciclo de vida: 4 semanas por ciclo sexual





Neurospora tipo selvagem (esquerda) e mutante (direita) crescendo em uma placa de Petri.

um único ascosporo para estudar. Assim, as análises nas quais ascos completos ou ascosporos aleatórios são usados são rápidas e diretas.

Como *Neurospora* é haplóide, fenótipos mutantes recém-obtidos são facilmente detectados com o uso de vários tipos de triagens e seleções. Um sistema favorito para o estudo do mecanismo de mutação é o gene *ad-3*, porque os mutantes *ad-3* são púrpura e facilmente detectados.

Embora os diplóides vegetativos de *Neurospora* não sejam prontamente obtidos, os geneticistas são capazes de criar um “diplóide mimético”, útil para testes de complementação e outras análises que requerem a presença de duas cópias de um gene (veja Capítulo 6). A fusão de duas linhagens diferentes produz um heterocácion, um indivíduo contendo dois tipos nucleares diferentes em um citoplasma comum. Os heterocárions também permitem o uso de uma versão do teste específico de locus, um modo de recuperar mutações em um alelo recessivo específico. (As células de heterocácion $+/m$ são plaqueadas e as colônias m/m são procuradas.)

Técnicas de Manipulação Genética

Mutagênese padrão:

Substâncias químicas e radiação	Mutações somáticas aleatórias
Mutagênese por transposon	Não-disponível

Transgênese:

Transformação mediada por plasmídeo	Inserção aleatória
-------------------------------------	--------------------

Nocautes gênicos direcionados:

RIP	Mutações GC \rightarrow AT em segmentos duplicados transgênicos antes de um cruzamento
Repressão	Inativação pós-transcricional somática de transgenes

Engenharia genética

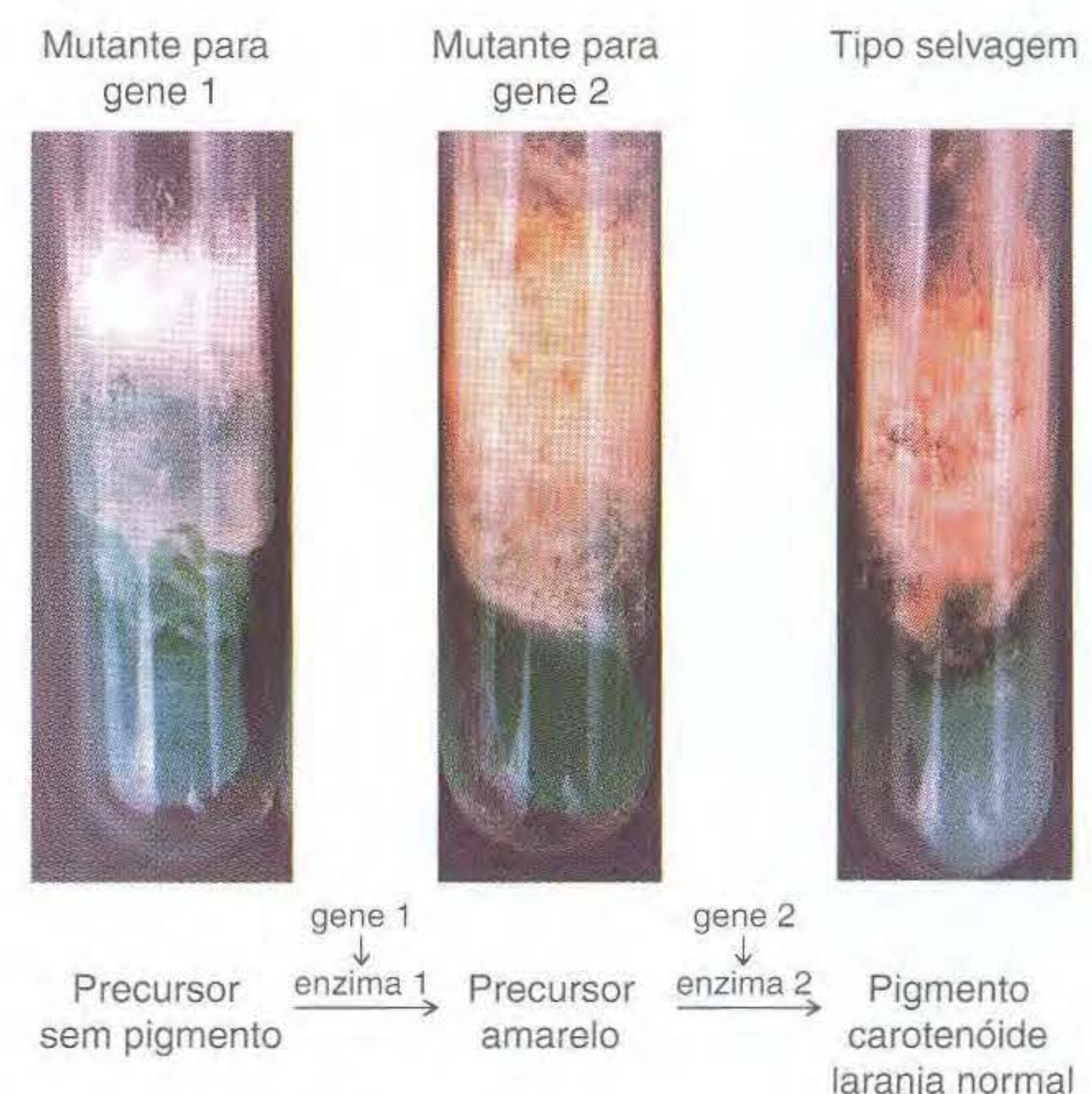
Transgênese. A primeira transformação eucariótica foi feita em *Neurospora*. Hoje em dia, a *Neurospora* é facilmente transformada com o uso de plasmídeos bacterianos que levam o transgene desejado, mais um marcador selecionável tal como resistência a higromicina para mostrar que o plasmídeo entrou. Nenhum plasmídeo replica-se em *Neurospora*, e, assim, um transgene é herdado apenas se for integrado a um cromossomo.

Nocautes direcionados. Em linhagens especiais de *Neurospora*, os transgenes freqüentemente se integram por recombinação homó-

loga. Assim, uma linhagem transgênica normalmente tem o gene residente mais o transgene homólogo, inserido em um local ectópico aleatório. Devido a essa duplicação de material, se a linhagem for cruzada, ela é sujeita a RIP, um processo genético que é único para *Neurospora*. RIP é um mecanismo pré-meiótico que introduz muitas transições GC para AT em ambas as cópias duplicadas, perturbando efetivamente o gene. RIP pode, portanto, ser usado como um modo conveniente para deliberadamente nocautear um gene específico.

Contribuições principais

George Beadle e Edward Tatum usaram *Neurospora* com organismo-modelo em seus estudos pioneiros sobre as relações gene-enzima, nas quais eles foram capazes de determinar as etapas enzimáticas na síntese de arginina (veja Capítulo 6). Seu trabalho com *Neurospora* estabeleceu o começo da genética molecular. Muitos estudos comparativos na genética do metabolismo celular com o uso de *Neurospora* se seguiram.



Via de síntese do pigmento carotenóide laranja em *Neurospora*.

Foi feito um trabalho pioneiro sobre genética dos processos meióticos, tais como crossing over e disjunção, e nos ritmos dos conídios. As culturas de crescimento contínuo mostram um ritmo diário de formação dos conidiosporos. Os resultados dos estudos pioneiros usando mutações que alteram esse ritmo contribuíram para um modelo geral para a genética dos ritmos circadianos.

Neurospora serve como modelo para os vários fungos filamentosos que afetam cultivos e os humanos, porque esses fungos em geral são difíceis de cultivar e manipular geneticamente. É ainda usado com um sistema de teste eucariótico simples para substâncias carcinogênicas e mutagênicas no ambiente humano.

Como podem ser feitos cruzamentos usando um genitor como fêmea, o ciclo é conveniente para o estudo de genética mitocondrial e interação núcleo-mitocôndrias. Uma ampla gama de plasmídeos mitocondriais lineares e circulares foram descobertos em isolados naturais. Alguns deles são retroelementos considerados intermediários na evolução dos vírus.

Outras áreas de contribuição

- Diversidade e adaptação de fungos
- Citogenética (base cromossômica da genética)
- Genes do tipo de reprodução
- Genes de compatibilidade de heterocácion (um modelo para a genética do reconhecimento do próprio e do alheio)



Arabidopsis thaliana

Organismo-chave para estudar:

- Desenvolvimento
- Expressão e regulação gênica
- Genômica de plantas



Arabidopsis thaliana, um membro da família de plantas Brassicaceae (couve), é relativamente recente como um organismo-modelo. A maior parte do trabalho foi feita nos últimos 20 anos. Não tem significado econômico: ela cresce prolificamente como uma praga em muitas partes temperadas do mundo. Entretanto, devido a seu pequeno tamanho, curto tempo de vida e genoma pequeno, ela superou os modelos genéticos de planta mais tradicionais, tais como o milho e o trigo, e tornou-se o modelo dominante para genética molecular de plantas.

Arabidopsis thaliana crescendo na natureza. As versões cultivadas em laboratório são menores. [Dan Tenaglia, www.missouriplants.com.]

"Estatísticas Vitais" Genéticas

Tamanho do genoma:	125 Mb
Cromossomos:	diplóide, 5 autossomos ($2n = 10$)
Número de genes:	25.000
Porcentagem com homólogos humanos:	18%
Tamanho médio dos genes:	2 kb, 4 íntrons/gene
Transposons:	10% do genoma
Genoma sequenciado em:	2000

Características especiais

Em comparação com outras plantas, *Arabidopsis* é pequena em relação tanto a seu tamanho físico quanto a seu genoma, características que são vantajosas para um organismo-modelo. *Arabidopsis* cresce a uma altura de menos de 10 cm sob condições apropriadas. Assim, ela pode crescer em grandes números, permitindo triagens em grande escala de mutantes e análises de prole. Seu tamanho genômico total de 125 Mb torna o genoma relativamente fácil de sequenciar comparado com outros genomas de organismos-modelo de plantas, tais como o genoma de milho (2.500 Mb) e o genoma de trigo (16.000 Mb).

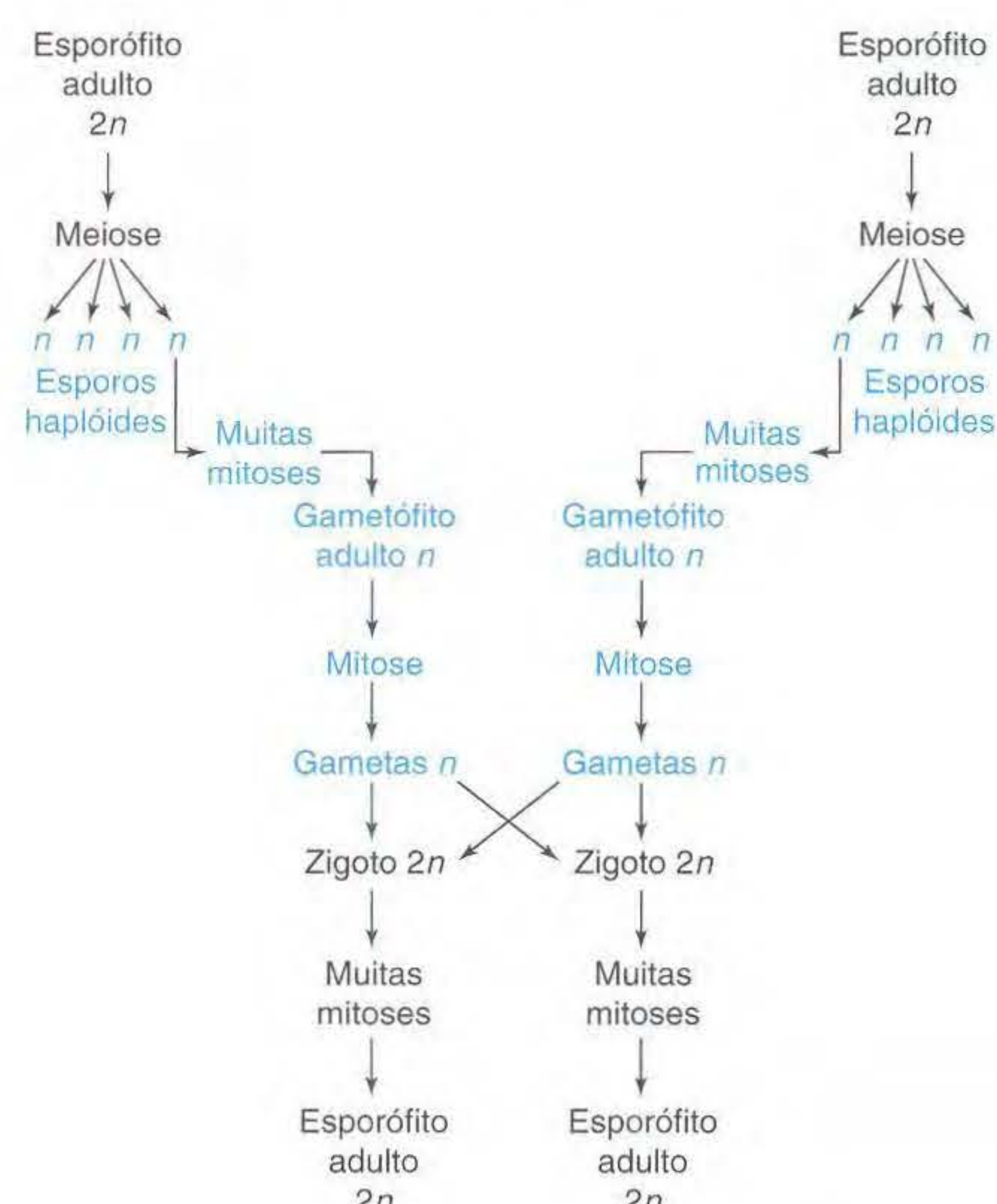
Análise genética

A análise de mutações de *Arabidopsis* pelos crossings baseia-se em métodos de ensaio e erro, essencialmente os usados por Mendel. Os estoques de plantas levando mutações úteis relevantes ao experimento são obtidos de centros públicos de estoque. As linhagens podem ser cruzadas manualmente umas com as outras ou autofecundadas. Embora as flores sejam pequenas, a polinização cruzada é facilmente feita removendo anteras que não abriram (que, às vezes, são comidas pelo experimentador como meio conveniente de descarte). Cada flor polinizada então produz uma longa vagem contendo um grande número de sementes. Essa produção abundante de prole (milhares de sementes por planta) é uma dádiva para os geneticistas que estão procurando mutantes raros ou outros eventos raros. Se uma planta leva uma nova mutação recessiva em sua linhagem germinativa, a autofecundação permite uma prole de homozigotos para a mutação recessiva ser recuperada nos descendentes imediatos da planta.

Ciclo de Vida

Arabidopsis tem um ciclo de vida familiar, com um estágio diplóide dominante. Uma planta possui várias flores, cada uma produzindo muitas sementes. Como muitas pragas anuais, seu ciclo de vida é rápido: leva apenas cerca de 6 semanas para que a semente de um plasmídeo produza uma nova quantidade de sementes.

Duração total do ciclo de vida: 6 semanas





Mutantes de *Arabidopsis*. (Esquerda) Flor tipo selvagem de *Arabidopsis*. (Meio) A mutação *agamous* (*ag*), que resulta em flores apenas com pétalas e sépalas (sem estruturas reprodutivas). (Direita) Um duplo mutante *ap1, cal*, que faz uma flor assemelhar-se a uma couve-flor. (Mutações similares em couve provavelmente são a causa das couves-flores.) [Fotos de George Haughn.]

Técnicas de Modificação Genética

Mutagênese padrão:

Substâncias químicas e radiação	Mutações aleatórias de linhagem germinativa ou somática
O próprio T-DNA ou transposons	Inserções marcadas aleatórias

Transgênese:

T-DNA leva o transgene	Inserção aleatória
------------------------	--------------------

Nocautes gênicos direcionados:

Mutagênese mediada por transposon ou T-DNA	Inserção aleatória; nocaute selecionado com PCR
RNAi	Mimetiza nocaute direcionado

Engenharia genética

Transgênese. O T-DNA de *Agrobacterium* é um vetor conveniente para introduzir transgenes (veja Capítulo 20). A construção vetor-transgene insere-se aleatoriamente no genoma. A transgênese oferece um modo efetivo de estudar a regulação gênica. O transgene é situado perto de um gene repórter tal como GUS, que produz um corante azul em qualquer posição na planta em que o gene é ativo.

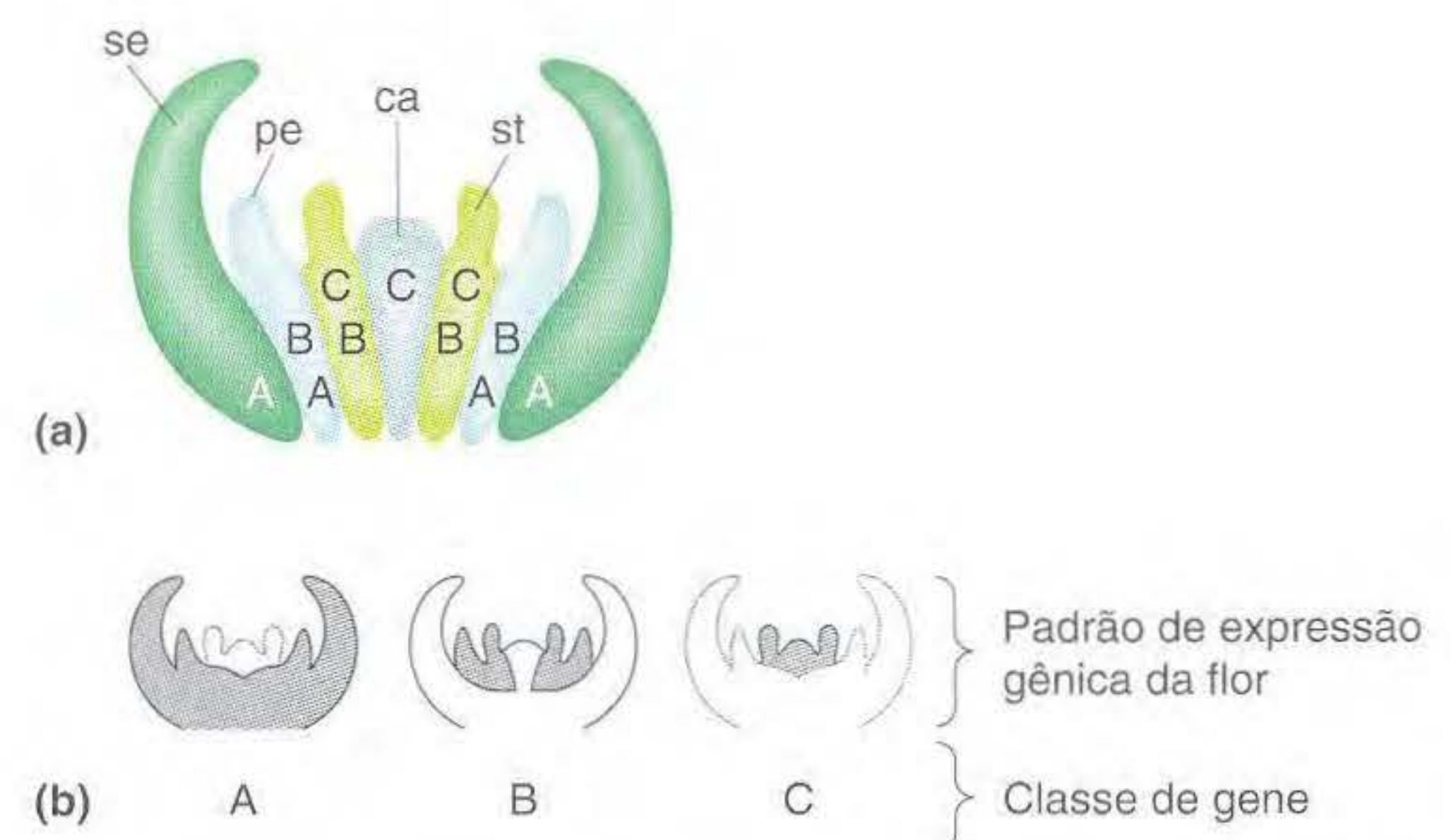
Nocautes direcionados. Como a recombinação homóloga é rara em *Arabidopsis*, genes específicos não podem ser facilmente nocauteados pela substituição homóloga com um transgene. Assim, em *Arabidopsis*, os genes são nocauteados pela inserção aleatória de um T-DNA vetor ou transposon (os transposons de milho tais como *Ac-Ds* são usados), e os nocautes gênicos específicos são selecionados aplicando-se análise de PCR de grandes conjuntos de plantas. A PCR usa uma sequência no T-DNA ou no transposon como um primer e uma sequência no gene de interesse como o outro primer. Assim, a PCR amplifica uma cópia do gene de interesse que leva uma inserção. Subdividindo-se o conjunto e repetindo-se o processo, leva-se à planta específica portadora do nocaute. Alternativamente, o RNAi também pode ser usado para inativar um gene específico.

Grandes coleções de mutantes de inserção de T-DNA estão disponíveis; eles têm as sequências flanqueadoras de plantas listadas em bancos de dados públicos; logo, se você está interessado em um gene específico, você pode ver se a coleção contém uma planta que tem uma inserção nesse gene. Uma característica conveniente das populações nocaute em plantas é que elas podem ser mantidas de modo fácil e barato como coleções de sementes por muitos anos

e, talvez, décadas. Essa característica não é possível para a maioria das populações de modelos animais. O verme *Caenorhabditis elegans* pode ser preservado como um animal congelado, mas as moscas-das-frutas (*Drosophila melanogaster*) não podem ser congeladas e recuperadas. Assim, linhagens de mutantes de moscas-das-frutas devem ser mantidas como organismos vivos.

Contribuições principais

Como o primeiro genoma de planta a ser seqüenciado, *Arabidopsis* forneceu um modelo importante para a arquitetura evolução do genoma de plantas. Além disso, os estudos de *Arabidopsis* fizeram contribuições importantes para nossa compreensão do controle genético do desenvolvimento de plantas. Os geneticistas isolaram mutações homeóticas que afetam o desenvolvimento de flores, por exemplo. Em tais mutantes, um tipo de parte floral é substituído por outro. A integração da ação desses mutantes levou a um modelo elegante de determinação da circularidade da flor com base em padrões de expressão de genes reguladores no meristema da flor. *Arabidopsis* também contribuiu amplamente para a base genética da fisiologia da planta, regulação gênica e interação das plantas e ambiente (incluindo a genética de resistência a doenças). Como *Arabidopsis* é uma planta natural da distribuição mundial, ela tem um grande potencial para o estudo da diversificação e adaptação evolutiva.



Estabelecimento do destino de um verticilo. (a) Os padrões de expressão gênica correspondentes aos diferentes destinos do verticilo. De fora para dentro, os destinos são sépala (se), pétala (pe), estame (st) e carpelo (ca). (b) As regiões sombreadas dos diagramas em corte do desenvolvimento da flor indicam os padrões de expressão gênica para os genes das classes A, B e C.

Outras áreas de contribuição

- Resposta ao estresse ambiental
- Sistemas de controle hormonal

Caenorhabditis elegans

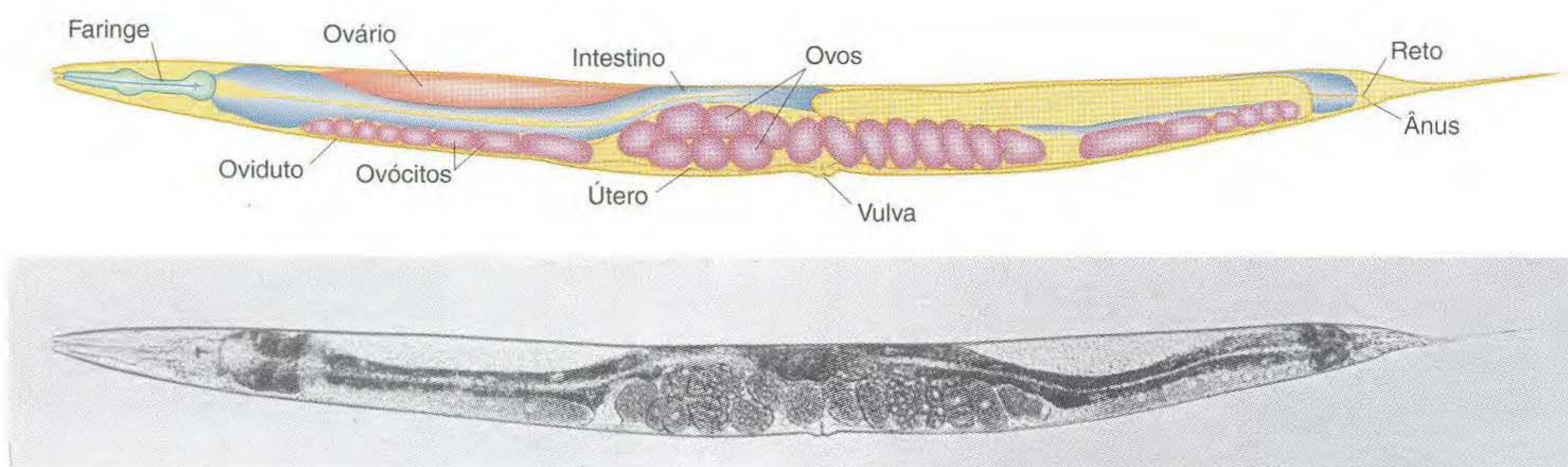
Organismo-chave para estudar:

- Desenvolvimento
- Comportamento
- Nervos e músculos
- Envelhecimento

"Estatísticas Vitais" Genéticas

Tamanho do genoma:	97 Mb
Cromossomos:	5 autossomos ($2n = 10$), cromossomo X
Número de genes:	19.000
Porcentagem com homólogos humanos:	25%
Tamanho médio dos genes:	5 kb, 5 éxons/gene
Transposons:	Vários tipos, ativos em algumas linhagens
Genoma seqüenciado em:	1998

Caenorhabditis elegans pode não aparecer muito ao microscópio, porém esse nematelminto de 1 mm de vida no solo é relativamente simples como animal. Mas essa simplicidade contribui para tornar *C. elegans* um bom organismo-modelo. Seu pequeno tamanho, crescimento rápido, habilidade de autofecundar-se, transparência e baixo número de células no corpo tornaram-no uma escolha ideal para o estudo da genética do desenvolvimento eucariótico.



Fotomicrografia e desenho de um *Caenorhabditis elegans* adulto.

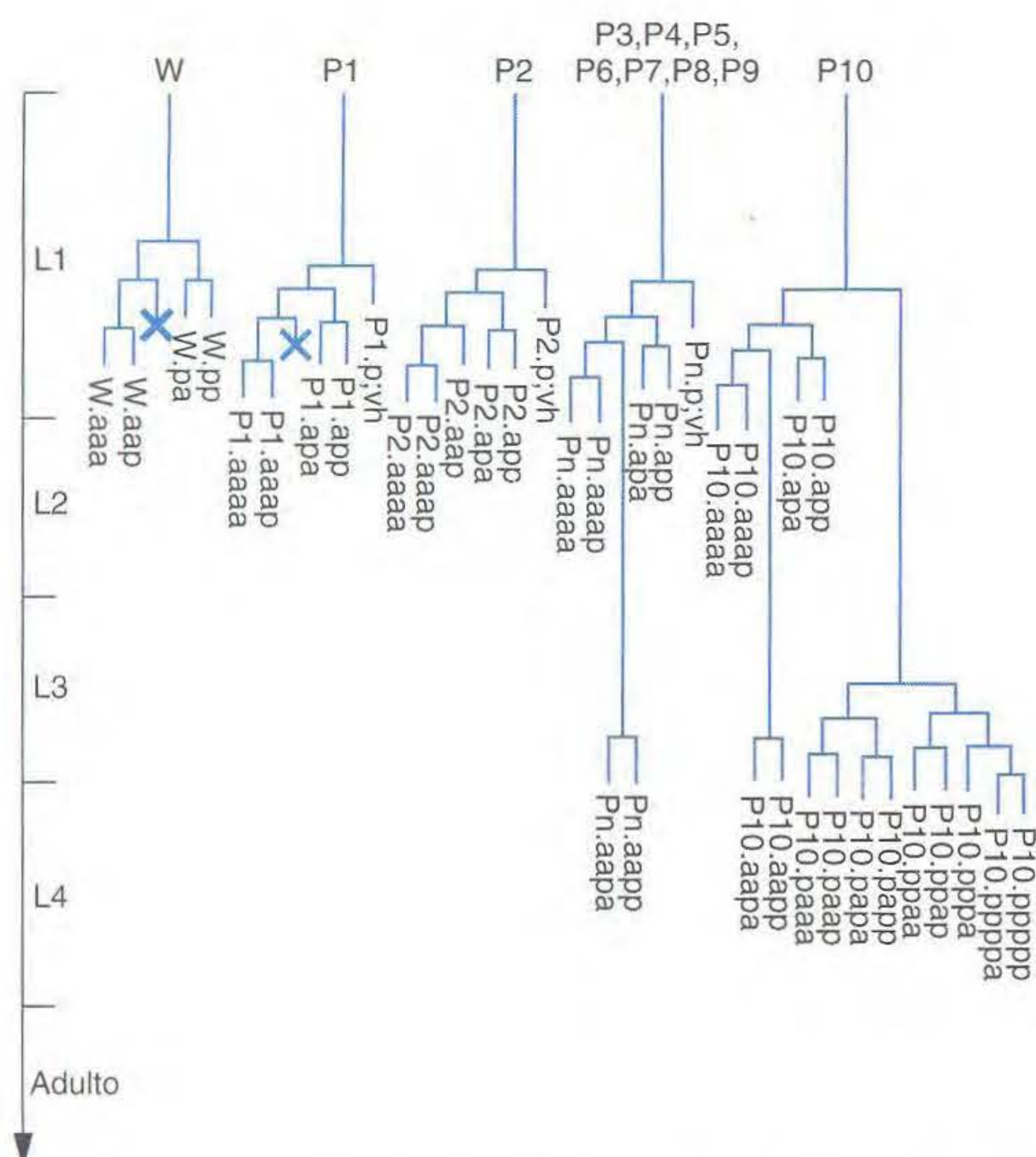
Características especiais

Os geneticistas podem ver através de *C. elegans*. Ao contrário de outros organismos-modelo multicelulares, tais como a mosca-das-frutas ou *Arabidopsis*, esse pequeno verme é transparente, o que o torna eficiente para triar grandes populações quanto a mutações interessantes que afetam praticamente qualquer aspecto da anatomia ou comportamento. A transparência também favorece os estudos do desenvolvimento: os pesquisadores podem acompanhar diretamente todos os estágios do desenvolvimento simples-

mente observando os vermes ao microscópio óptico. Os resultados de tais estudos mostraram que o desenvolvimento de *C. elegans* é rigidamente programado, e que cada verme tem um número surpreendentemente pequeno e consistente de células (959 nos hermafroditas e 1.031 nos machos). De fato, os biólogos acompanharam os destinos de células específicas à medida que o verme se desenvolvia e determinaram o padrão exato da divisão celular que leva a cada órgão adulto. Esse esforço produziu um pedigree de linhagem de cada célula adulta (veja Capítulo 12).

Análise Genética

Como os vermes são pequenos e se reproduzem rápida e prolificamente (a autofecundação produz cerca de 300 na prole e o cruzamento cerca de 1.000), eles produzem grandes proles que podem



Uma representação simbólica das linhagens de 11 células. Uma célula que sofre morte celular programada é indicada por X azul ao final da ramificação de uma linhagem.

Ciclo de Vida

C. elegans é único entre os principais modelos animais porque um dos dois sexos é hermafrodita (XX). O outro é macho (XO). Os dois sexos podem ser diferenciados pelo maior tamanho dos hermafroditas e pelas diferenças em seus órgãos sexuais. Os hermafroditas produzem tanto ovócitos quanto espermatozóides, e, assim, eles podem se autofecundar. A prole de um hermafrodita autofecundado também é de hermafroditas, exceto quando uma rara não-disjunção leva a um macho XO. Se hermafroditas e machos são misturados, os sexos copulam, e muitos dos zigotos resultantes terão sido fertilizados pelos espermatozóides amebóides do macho. A fertilização e o embrião produzido ocorrem dentro do hermafrodita, que então põe ovos. Os ovos terminam seu desenvolvimento externamente.

Duração total do ciclo de vida: 3 dias e $\frac{1}{2}$

ser triadas quanto a eventos genéticos raros. Além disso, como o hermafroditismo em *C. elegans* possibilita a autofecundação, vermes individuais com mutações recessivas homozigotas podem ser recuperados rapidamente por autofecundação da prole de vermes individuais tratados. Ao contrário, outros modelos animais, tais como a mosca-das-frutas ou camundongos, requerem reproduções entre irmãos e levam muito mais gerações para recuperar mutações.

Técnicas de Modificação Genética

Mutagênese padrão:

Substâncias químicas (EMS) e radiação	Mutações aleatórias na linhagem germinativa
Transposons	Inserções aleatórias na linhagem germinativa

Transgênese:

Injeção transgênica de gônada	Arranjos transgênicos não-integrados; integração ocasional
-------------------------------	--

Nocautes gênicos direcionados:

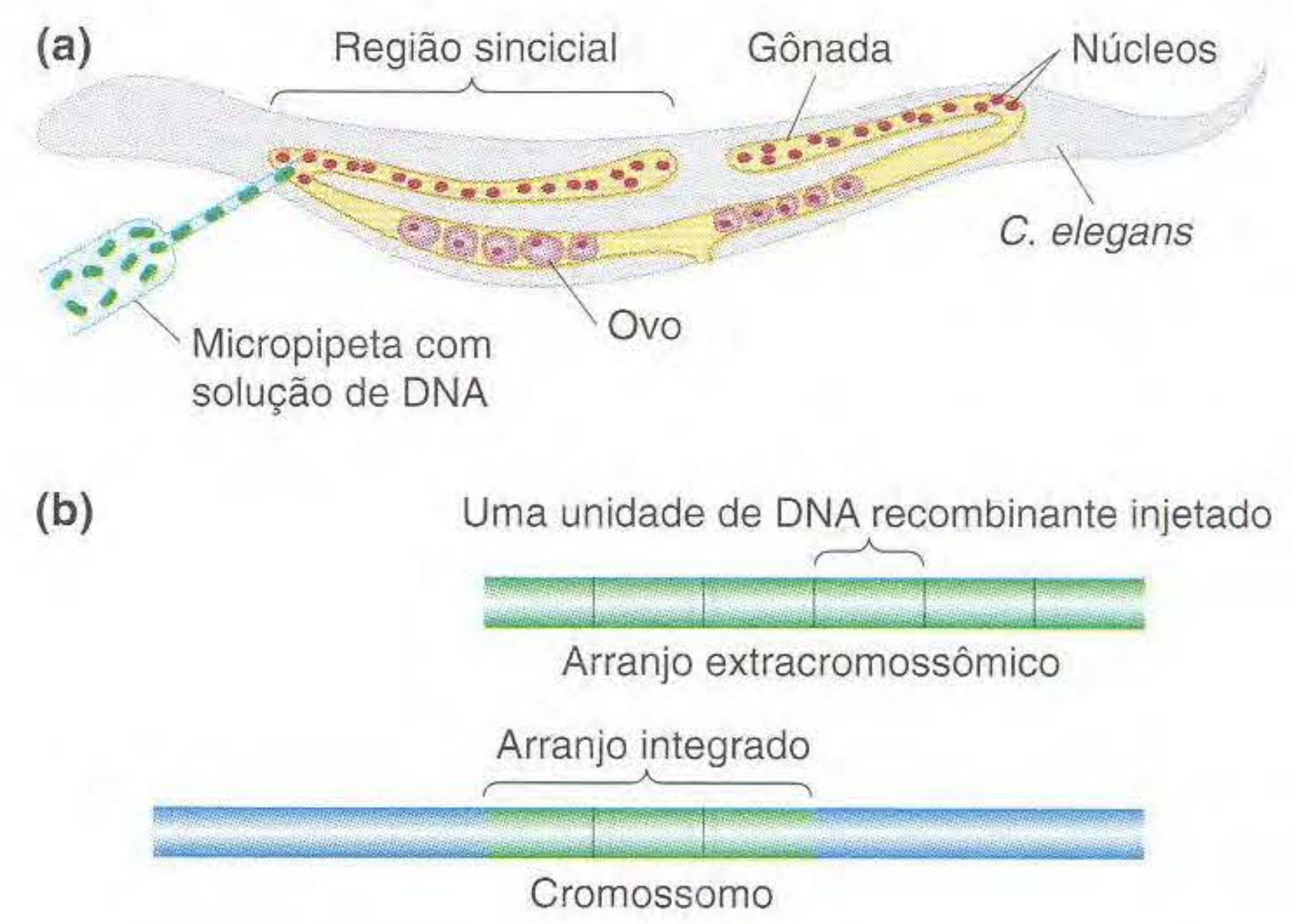
Mutagênese mediada por transposons	Nocautes selecionados com PCR
RNAi	Mimetiza nocautes direcionados
Ablação com laser	Nocaute de uma célula

Engenharia genética

Transgênese. A introdução de transgenes na linhagem germinativa é possibilitada por uma propriedade especial das gônadas de *C. elegans*. As gônadas do verme são sinciciais, significando que existem muitos núcleos em um citoplasma comum. Os núcleos não se tornam incorporados às células até a meiose, quando começa a formação de ovócitos individuais ou espermatozoides. Assim, uma solução de DNA contendo o transgene injetado na gônada de um hermafrodita expõe mais de 100 núcleos precursores de células germinativas ao transgene. Ao acaso, alguns desses núcleos irão incorporar-se ao DNA (veja Capítulo 20).

Os transgenes recombinam-se para formar arranjos multicópia em tandem. Em um ovócito, os arranjos não se integram a um cromossomo, mas os transgenes dos arranjos ainda são expressos. Assim, o gene levado em um clone de DNA tipo selvagem pode ser identificado se introduzido em uma linhagem receptora recessiva específica (complementação funcional). Em alguns casos, mas não em todos, os arranjos transgênicos são passados adiante para a prole. Para aumentar a chance de herança, os vermes são expostos a radiação ionizante, que pode induzir a integração de um arranjo em uma posição cromossômica ectópica, e, nesse local, o arranjo é confiavelmente transmitido para a prole.

Nocautes direcionados. Nas linhagens com transposons ativos, os próprios transposons tornam-se agentes de mutação inserindo-se em locais aleatórios no genoma, nocauteando os genes interrompidos. Se pudermos identificar os organismos com inserções em um gene específico de interesse, poderemos isolar um nocaute gênico direcionado. As inserções em genes específicos podem ser detectadas usando PCR se um primer de PCR for baseado na sequência do transposon e outro baseado na sequência do gene de interesse. Alternativamente, o RNAi pode ser usado para anular a função de genes específicos. Como uma alternativa para mutação, as células individuais podem ser mortas por um feixe de laser para observar o efeito no funcionamento ou desenvolvimento do verme.



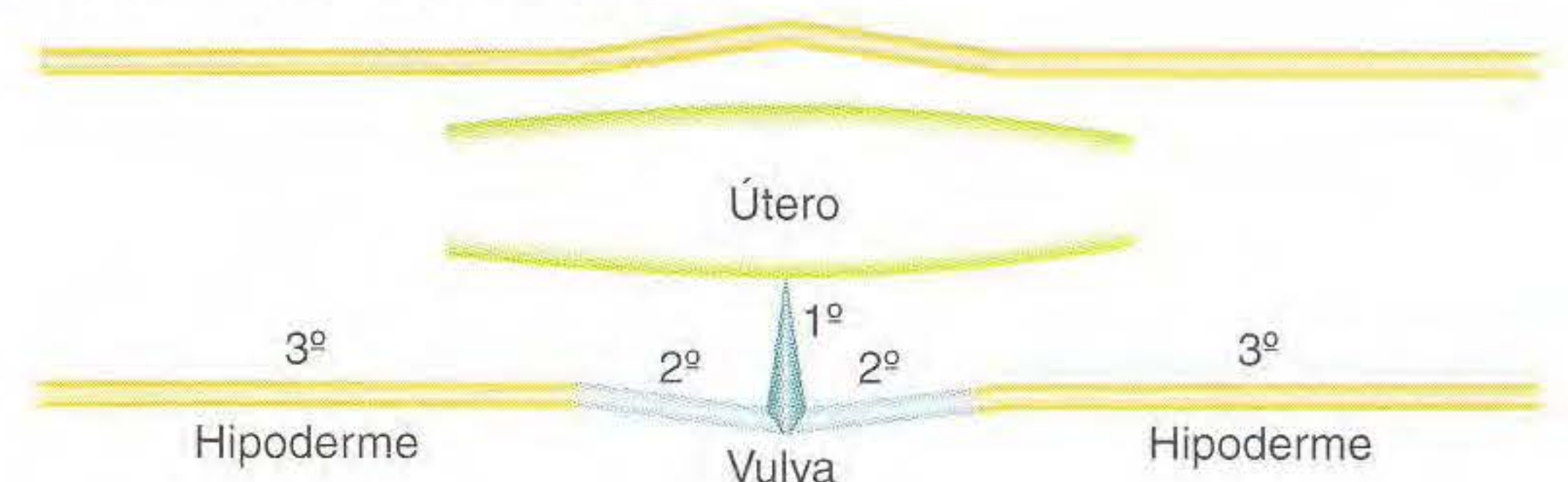
Criação de transgenes de *C. elegans*. (a) Método de injeção. (b) Arranjos extracromossômicos e integrados.

Contribuições principais

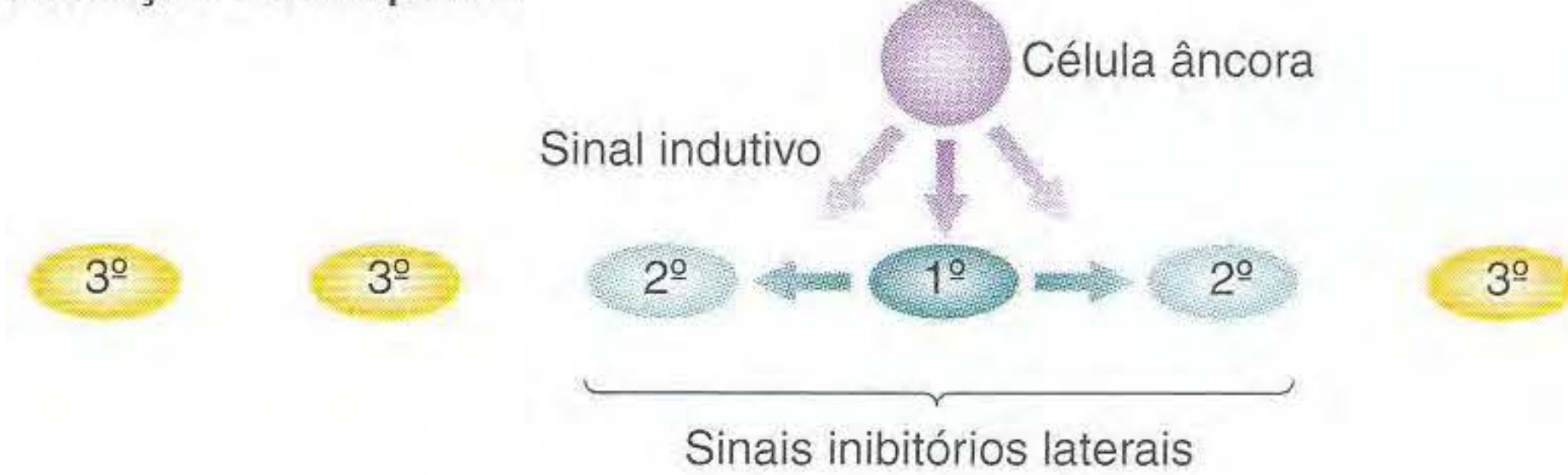
C. elegans tornou-se o organismo-modelo favorito para o estudo de vários aspectos do desenvolvimento devido a seu pequeno e invariável número de células. Um exemplo é a morte celular programada, um aspecto crucial do desenvolvimento normal. Algumas células são geneticamente programadas para morrer no curso do desenvolvimento (um processo chamado de apoptose). Os resultados dos estudos de *C. elegans* contribuíram para um modelo geral útil de apoptose, que também é conhecido como uma característica do desenvolvimento humano.

Outro sistema-modelo é o desenvolvimento da vulva, a abertura para o exterior da via reprodutiva. Os hermafroditas com vulvas defeituosas darão uma prole que, nas triagens, são facilmente visíveis agrupadas dentro do corpo. Os resultados dos estudos de hermafroditas sem vulva ou com muitas vulvas revelaram como as células completamente equivalentes podem tornar-se diferenciadas em tipos celulares diferentes (veja Capítulo 12).

(a) Tecido derivado de 1ª, 2ª e 3ª, células



(b) Sinalização subsequente



Produção da vulva de *C. elegans*. (a) O tecido final diferenciado. (b) Método de diferenciação. As células começam completamente equivalentes. Uma célula âncora atrás das células equivalentes manda um sinal para as células vizinhas, que se tornam a vulva. A célula vulvar primária então manda um sinal lateral para suas vizinhas, impedindo que elas se tornem células primárias, muito embora elas também tenham recebido o sinal da célula vulvar.

O comportamento também foi objeto de dissecação genética. *C. elegans* oferece uma vantagem de que os vermes com comportamento defeituoso podem geralmente viver e se reproduzir. Os sistemas nervoso e muscular do verme foram geneticamente dissecados, permitindo que os comportamentos fossem ligados a genes específicos.

Outras áreas de contribuição

- Sinalização célula-célula



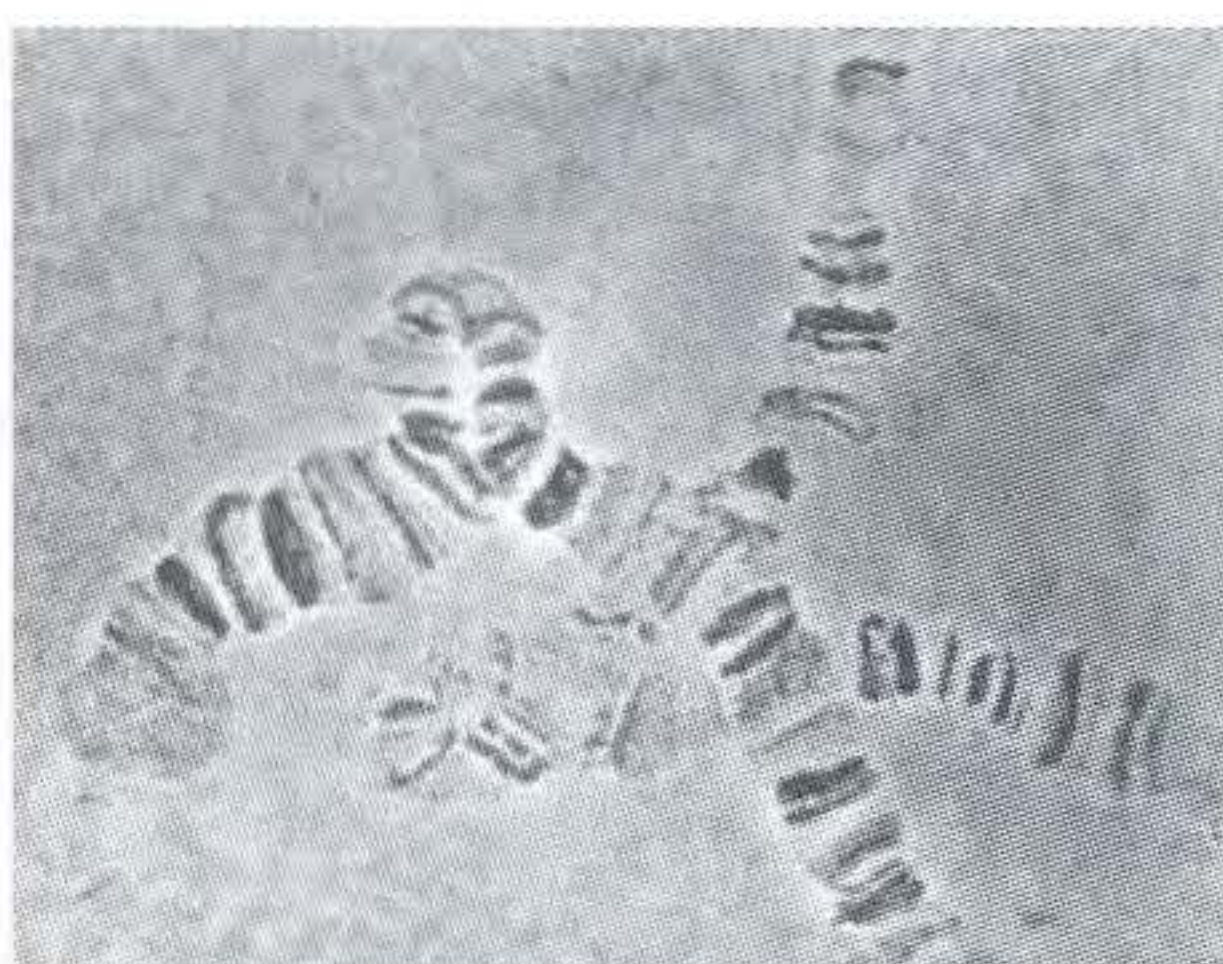
Drosophila melanogaster

Organismo-chave para estudar:

- Genética de transmissão
- Citogenética
- Desenvolvimento
- Genética de populações
- Evolução

"Estatísticas Vitais" Genéticas

Tamanho do genoma:	180 Mb
Cromossomos:	Diplóide, 3 autossomos, X e Y ($2n = 8$)
Número de genes:	13.000
Percentagem com homólogos humanos:	~ 50%
Tamanho médio dos genes:	3 kb, 4 éxons/gene
Transposons:	Elementos P, entre outros
Genoma sequenciado em:	2000



Cromossomos politênicos.

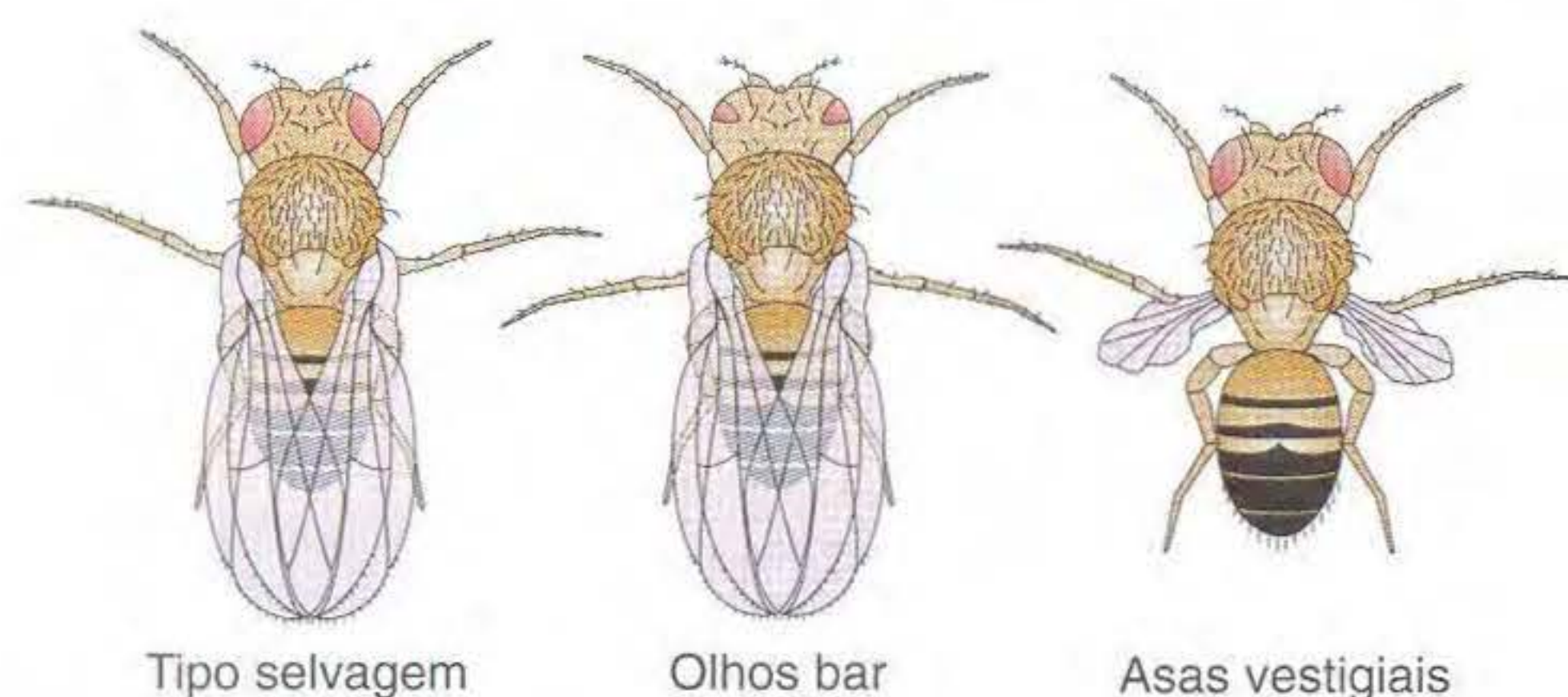
A mosca-das-frutas *Drosophila melanogaster* (imprecisamente traduzida como "atraída por fermentados") foi um dos primeiros organismos-modelo usados em Genética. Ela foi escolhida em parte porque está prontamente disponível em frutos maduros, tem um curto ciclo vital do tipo diplóide e é simples de cultivar e cruzar em vidros contendo uma camada de alimentos. As primeiras análises genéticas mostraram que seus mecanismos de herança são fortemente similares aos de outros eucariontes, destacando seu papel como um organismo-modelo. Sua popularidade como organismo-modelo entrou em declínio durante os anos quando *E. coli*, leveduras e outros microrganismos estavam sendo desenvolvidos como ferramentas moleculares. Entretanto, a *Drosophila* teve um renascimento porque se presta muito bem ao estudo da base genética do desenvolvimento, uma das questões centrais da Biologia. A importância da *Drosophila* como um modelo para genética humana é demonstrada pela descoberta de que aproximadamente 60% dos genes conhecidos causadores de doença em humanos, bem como 70% dos genes de câncer, têm contrapartes na *Drosophila*.

Características especiais

A *Drosophila* tornou-se popular como um organismo experimental no início do século XX devido a características comuns à maioria dos organismos-modelo. Ela é pequena (3 mm de tamanho), simples de criar (originalmente, em garrafas de leite), com reprodução rápida (apenas 12 dias do zigoto ao adulto) e fácil de obter (é só ter uma fruta apodrecendo). Foi demonstrado que ela tem uma grande gama de alelos mutantes interessantes que foram usados para estabelecer as regras básicas da genética de transmissão. Os primeiros pesquisadores também tiraram proveito de uma característica única da mosca-das-frutas: os cromossomos politênicos (veja Capítulo 16). Nas glândulas salivares e em alguns outros tecidos, esses "cromossomos gigantes" são produzidos por várias rodadas de replicação de DNA sem segregação cromossômica. Cada cromossomo politênico apresenta um único padrão de bandeamento, dando aos geneticistas marcos que podem ser usados para correlacionar mapas baseados em recombinação com cromossomos. Esses avanços iniciais, juntamente com a grande quantidade de conhecimentos acumulados sobre o organismo, tornaram a *Drosophila* um modelo genético atraente.

Análise genética

Os cruzamentos em *Drosophila* podem ser feitos bem facilmente. Os genitores podem ser tipo selvagem ou estoques de mutantes obtidos de centros de estoque ou como novas linhagens mutantes.

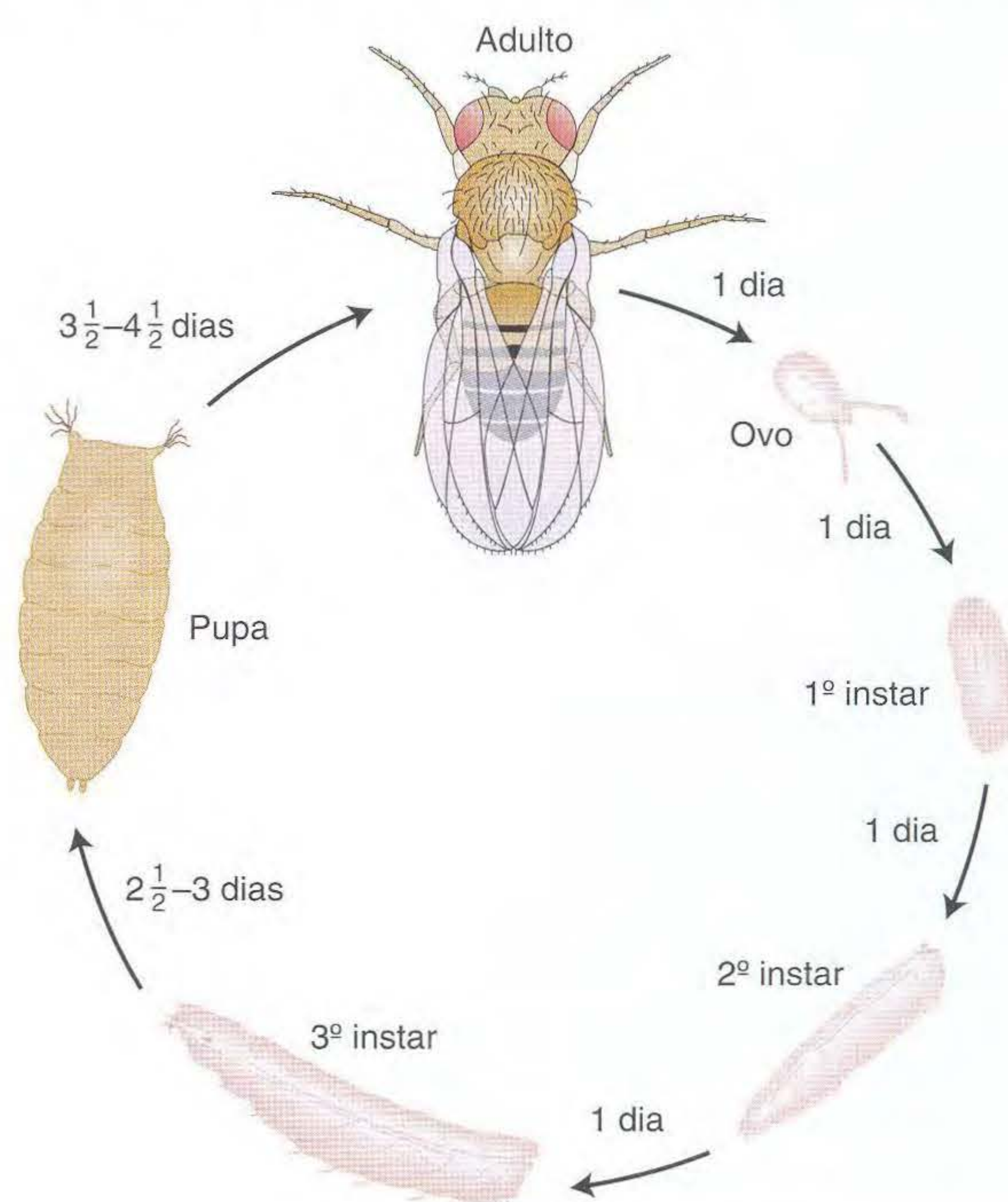


Dois mutantes morfológicos de *Drosophila*, com o tipo selvagem para comparação.

Ciclo de Vida

A *Drosophila* tem um curto ciclo de vida diplóide que se presta bem à análise genética. Após eclodir de um ovo, a mosca desenvolve-se por vários estágios larvais e um estágio de pupa antes de emergir como um adulto, que logo se torna sexualmente maduro. O sexo é determinado pelos cromossomos X e Y (XX é fêmea, XY é macho), embora, ao contrário dos humanos, o número de X em relação ao número de autossomos determina o sexo (veja Capítulo 2).

Duração total do ciclo de vida: 12 dias do ovo ao adulto



Para fazer um cruzamento, machos e fêmeas são colocados juntos em uma garrafa, e as fêmeas põem ovos em um alimento semi-sólido que cobre o fundo da garrafa. Após surgirem as pupas, a prole pode ser anestesiada para a contagem dos membros de classes fenotípicas e distinguir os machos das fêmeas (por seus padrões diferentes de faixas abdominais). Entretanto, como a prole feminina permanece virgem por apenas algumas horas após o surgimento das pupas, elas devem ser imediatamente isoladas se forem usadas para fazer cruzamentos controlados. Os cruzamentos destinados a fazer combinações gênicas específicas devem ser cuidadosamente planejados, pois não ocorre crossing nos machos de *Drosophila*. Assim, no macho, os genes ligados não se recombina para ajudar a criar novas combinações.

Para obter novas mutações recessivas, programas especiais de cruzamento (dos quais o protótipo é o teste CLB de Muller) fornecem sistemas convenientes de triagem. Nesses testes, as moscas mutagenizadas são cruzadas com um estoque tendo um cromossomo balanceador (veja Capítulo 16). As mutações recessivas são tornadas homozigotas por endocruzamento por uma ou duas gerações, começando com moscas F_1 isoladas.

Técnicas de Modificação Genética

Mutagênese padrão:

Substâncias químicas (EMS) e radiação	Mutações somáticas e germinativas aleatórias
---------------------------------------	--

Transgênese:

Mediada por elemento <i>P</i>	Inserção aleatória
-------------------------------	--------------------

Nocautes gênicos direcionados:

Substituição induzida	Sai alelo ectópico nulo e recombina-se com alelo tipo selvagem
RNAi	Mimetiza nocaute direcionado

Engenharia genética

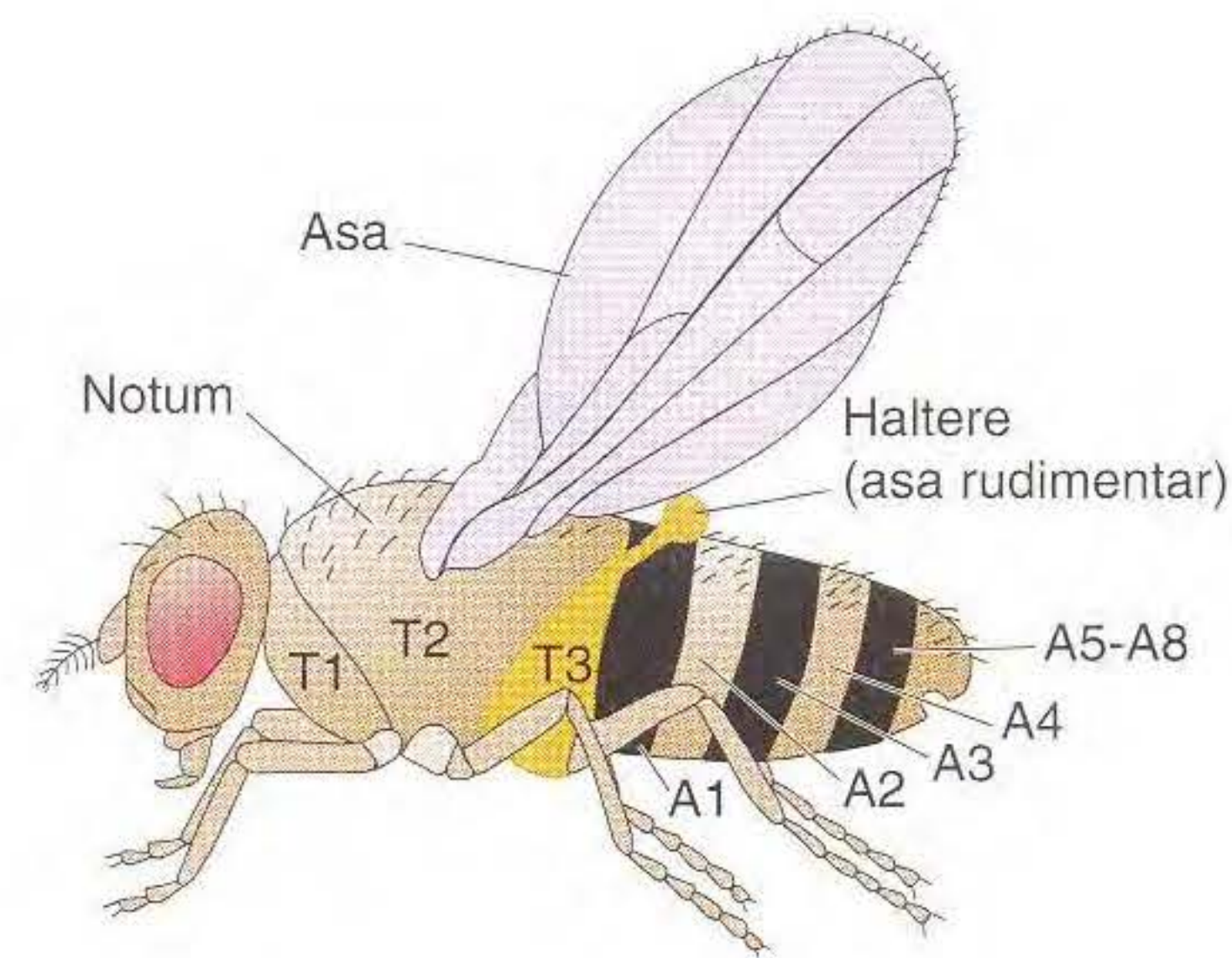
Transgênese. A construção de moscas transgênicas requer a ajuda de um transposon de *Drosophila* chamado de elemento *P*. Os geneticistas constroem um vetor que leva um transgene flanqueado por repetições de elemento *P*. O vetor do transgene é então injetado em um zigoto juntamente com um plasmídeo auxiliar contendo uma transposase. A transposase permite que o transgene salte aleatoriamente no genoma nas células germinativas do embrião (veja Capítulo 20).

Nocautes direcionados. Os nocautes gênicos direcionados podem ser feitos, primeiro, introduzindo-se um alelo nulo transgenicamente em uma posição ectópica e, segundo, induzindo-se enzimas especiais que causam a remoção do alelo nulo. O fragmento removido (que é linear) então encontra e substitui a cópia endógena por crossing over homólogo. Entretanto, os nocautes funcionais podem ser produzidos mais eficientemente por RNAi.

Contribuições principais

Grande parte do desenvolvimento inicial da teoria cromossômica da hereditariedade foi baseada nos resultados dos estudos de *Drosophila*. Os geneticistas trabalhando com *Drosophila* fizeram avanços importantes no desenvolvimento de técnicas para mapeamento

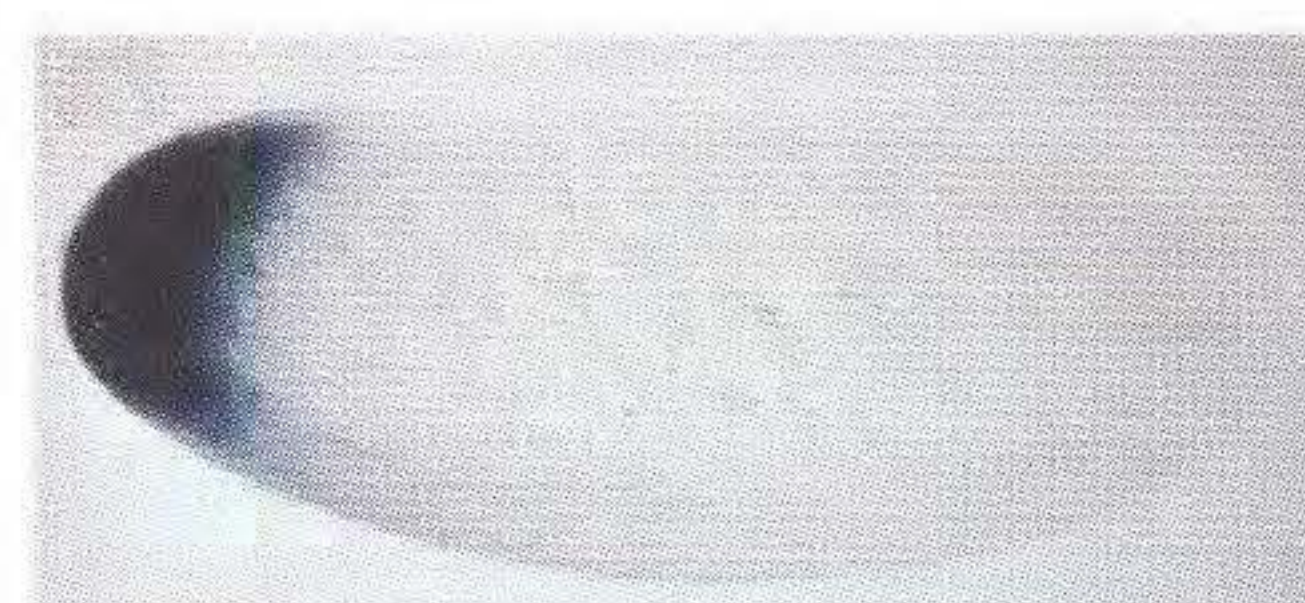
Os segmentos torácicos e abdominais normais de *Drosophila*.



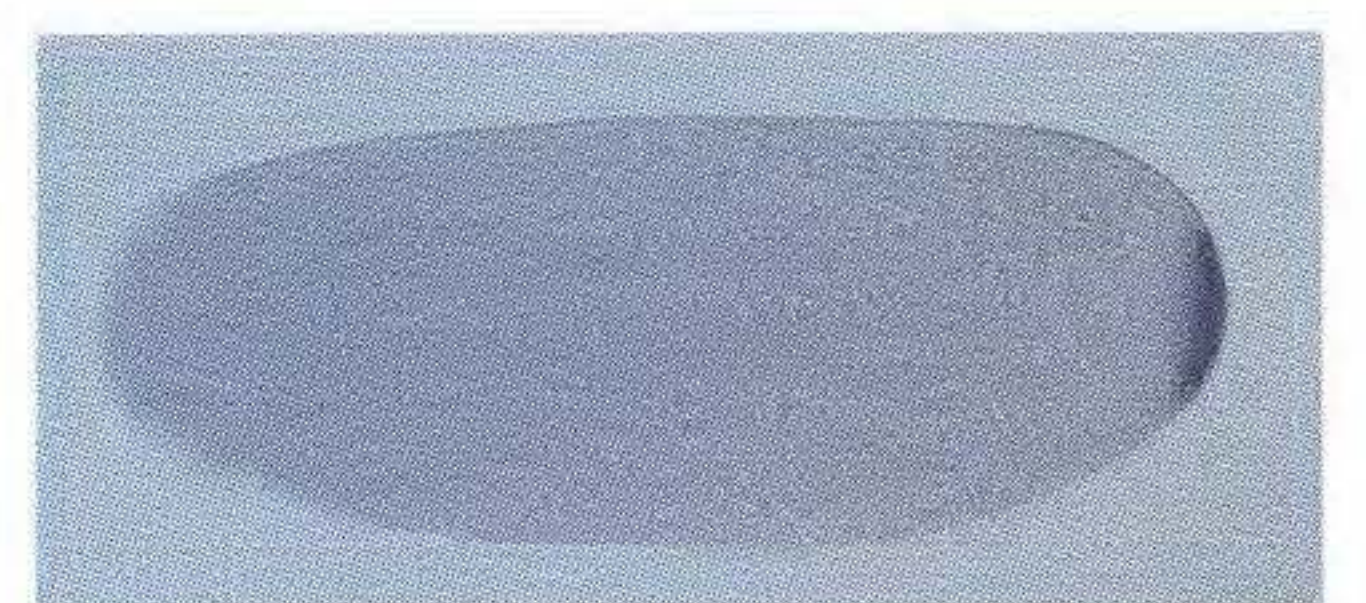
gênico, na compreensão da origem e natureza da mutação gênica e na documentação da natureza e comportamento dos rearranjos cromossômicos (veja Capítulo 4).

Suas descobertas abriram as portas para outros estudos pioneiros:

- Os primeiros estudos de cinética da indução de mutação e medida das taxas de mutação foram feitos com o uso de *Drosophila*. O teste CLB de Muller e outros similares deram métodos convenientes de triagem para mutações recessivas.
- Os rearranjos cromossômicos que movem genes adjacentes para heterocromatina foram usados para descobrir e estudar variação de efeito de posição.
- Na última parte do século XX, após a identificação de algumas classes importantes de mutações, tais como a homeótica e as mutações de efeito materno, a *Drosophila* assumiu um papel central na genética do desenvolvimento, um papel que continua até hoje (veja Capítulo 12). As mutações de efeito materno que afetam o desenvolvimento dos embriões, por exemplo, foram cruciais na elucidação das determinações genéticas do plano corpóreo em *Drosophila*; essas mutações são identificadas pela triagem do desenvolvimento de fenótipos anormais nos embriões de uma fêmea específica. As técnicas tais como as triagens de captura de acentuador permitiram a descoberta de novas regiões regulatórias no genoma que afetam o desenvolvimento. Por esses métodos e outros, os biólogos de *Drosophila* fizeram avanços importantes na compreensão e determinação da segmentação e dos eixos corporais. Algumas das principais descobertas de genes, tais como genes homeóticos, têm ampla relevância em animais em geral.



(a) mRNA *Bicoid*



(b) mRNA *nos*

Fotomicrografias mostrando gradientes de determinantes de eixo corporal. (a) mRNA para o gene *bcd* é mostrado localizado na ponta anterior (esquerda) do embrião. (b) mRNA do gene *nos* está situado no lado posterior (à direita da ponta do embrião). A distribuição das proteínas codificadas por esses genes e outros genes determina o eixo corpóreo.

Outras áreas de contribuição

- Genética de populações
- Genética evolutiva
- Genética do comportamento



Mus musculus

Organismo-chave para estudar:

- Doenças humanas
- Mutação
- Desenvolvimento
- Cor da pelagem
- Imunologia

"Estatísticas Vitais" Genéticas

Tamanho do genoma:	2.600 Mb
Cromossomos:	19 autossomos, X e Y ($2n = 40$)
Número de genes:	30.000
Porcentagem com homólogos humanos:	99%
Tamanho médio dos genes:	40 kb, 8,3 éxons/gene
Transposons:	Fonte de 38% do genoma
Genoma sequenciado em:	2002



Como os humanos e a maioria dos animais domesticados são mamíferos, a genética dos mamíferos é de grande interesse para nós. Entretanto, os mamíferos não são ideais para a genética: eles são relativamente grandes em tamanho comparados com outros organismos-modelo, necessitando assim de grandes e caras instalações; seus ciclos de vida são longos, e seus genomas, grandes e complexos. Comparados com outros mamíferos, entretanto, os camundongos (*Mus musculus*) são relativamente pequenos, têm ciclos de vida curtos e são facilmente obtidos, o que os torna uma excelente escolha para um modelo de mamífero. Além disso, os camundongos têm uma vantagem em genética porque "peculiaridades" já foram desenvolvidas em muitas linhagens de interesse em camundongos que fornecem uma fonte de variantes para a análise genética. As pesquisas de genética mendeliana de camundongos começaram no início do século XX.

Um camundongo adulto e sua prole.

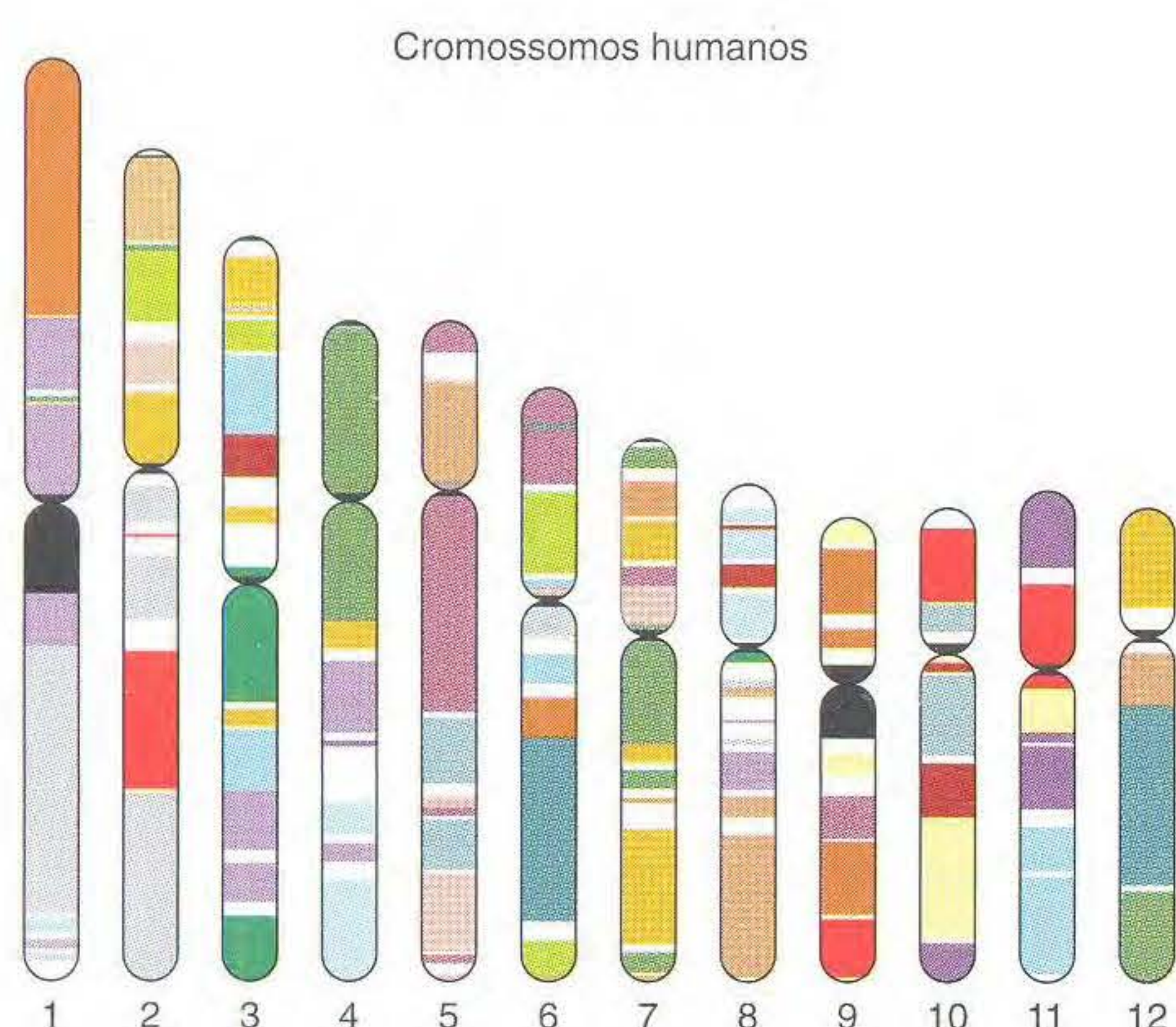
Características especiais

Os camundongos não são exatamente pequenos humanos peludos, mas sua constituição genética é muito similar à nossa. Entre os organismos-modelo, o camundongo é aquele cujo genoma mais proximamente assemelha-se ao genoma humano. O genoma do camundongo é cerca de 14% menor que o dos humanos (o genoma humano tem 3.000 Mb), mas tem aproximadamente o mesmo número de genes (as estimativas atuais são de um pouco menos de 30.000). Surpreendentemente, 99% dos genes de camundongo parecem ter homólogos em humanos. Além disso, em grande proporção, o genoma é sintênico com o de humanos; isto é, existem grandes blocos contendo os mesmos genes nas mesmas posições relativas (veja Capítulos 13 e 19). Tais similaridades genéticas são a chave para o sucesso dos camundongos como organismo-modelo; essas similaridades permitem que o camundongo seja tratado como

"preferido" por suas contrapartes humanas de muitos modos. Mutágenos e carcinógenos potenciais que suspeitamos sejam causadores de danos em humanos, por exemplo, são testados em camundongos, e os modelos de camundongo são essenciais no estudo de uma ampla gama de doenças genéticas humanas.

Análise genética

Camundongos mutantes e "tipo selvagem" (embora não sejam de fato da selva) são fáceis de obter: eles podem ser solicitados a centros com grandes estoques que fornecem camundongos adequados a cruzamentos de vários outros tipos de experimentos. Muitas dessas linhagens são derivadas de camundongos cruzados em séculos passados por criadores. Os cruzamentos controlados podem ser feitos simplesmente pareando um macho com uma fêmea não-grávida. Na maioria dos casos, os genótipos parentais podem ser dados pelo macho ou pela fêmea.

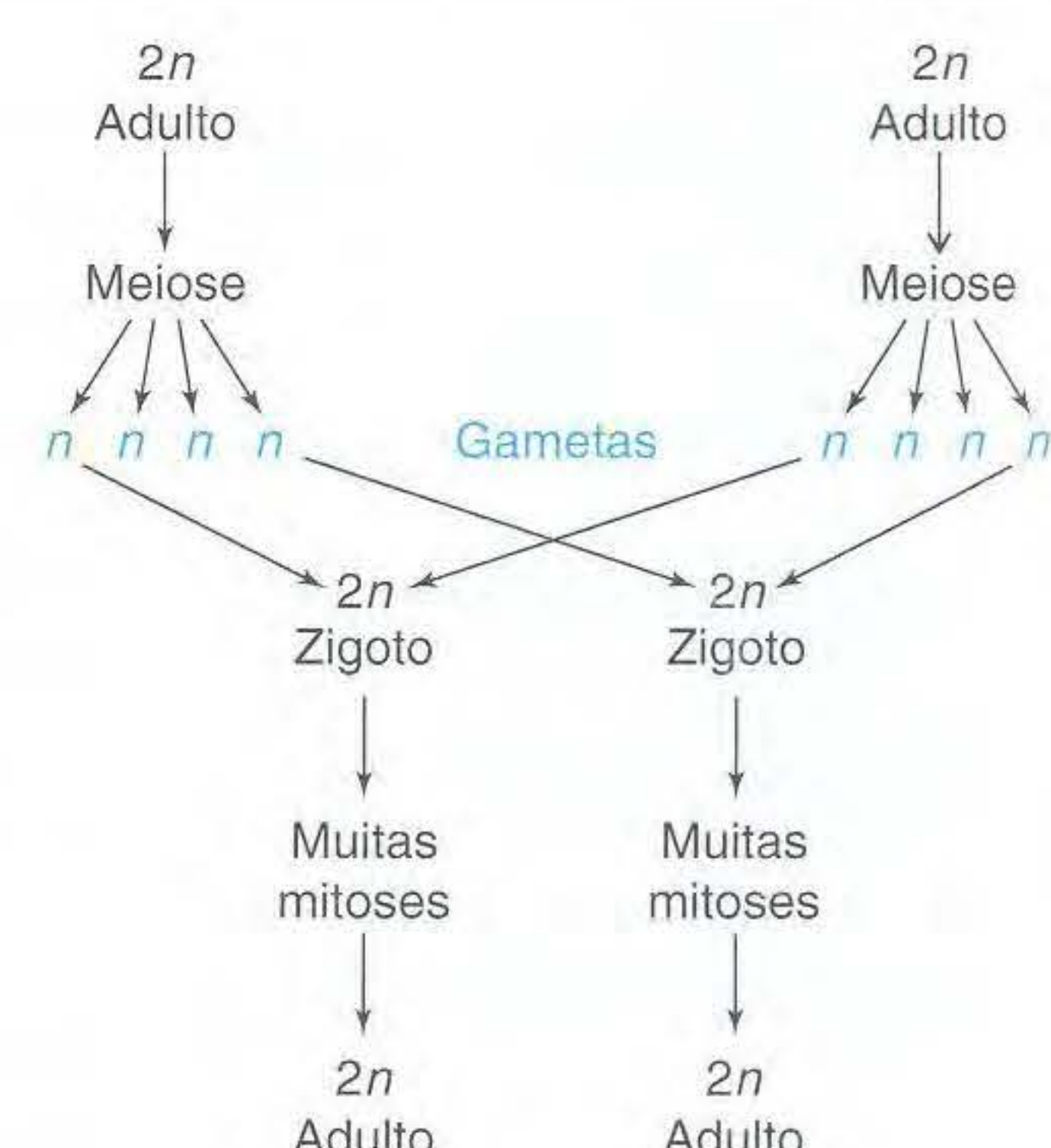


Um mapa de sintenia camundongo-humano dos 12 cromossomos do genoma humano. O código de cores é usado para mostrar as correspondências regionais de cada bloco do genoma humano com as seções correspondentes do genoma humano. Cada cor representa um cromossomo diferente de camundongo.

Ciclo de Vida

Os camundongos têm um ciclo de vida diplóide familiar, com um sistema XY de determinação de sexo similar ao dos humanos. As ninhadas são de 5 a 10 filhotes; entretanto, a fecundidade das fêmeas declina após cerca de 9 meses e, assim, raramente têm mais de cinco ninhadas.

Duração total do ciclo de vida: 10 semanas do nascimento até ter filhotes, na maioria das linhagens laboratoriais



A maioria das estimativas-padrão das taxas de mutação de mamíferos (incluindo os humanos) são estimadas em camundongos. De fato, os camundongos fornecem o teste final de agentes suspeitos de causar mutações em humanos. As taxas de mutação na linhagem germinativa são medidas com o uso de teste específico de locus: mutagenizar gônadas $+/+$, cruzar com m/m (m é uma mutação recessiva conhecida no locus em estudo) e procurar prole m^*/m (m^* é uma nova mutação). O procedimento é repetido para sete loci de amostra. A medida das taxas de mutação somáticas usa um modo similar, mas o mutágeno é injetado no feto. Os camundongos foram amplamente usados para o estudo do tipo de mutação somática que dá origem ao câncer.

Técnicas de Modificação Genética

Mutagênese padrão:

Substâncias químicas e radiação Mutações somáticas e de linhagem germinativa

Transgênese:

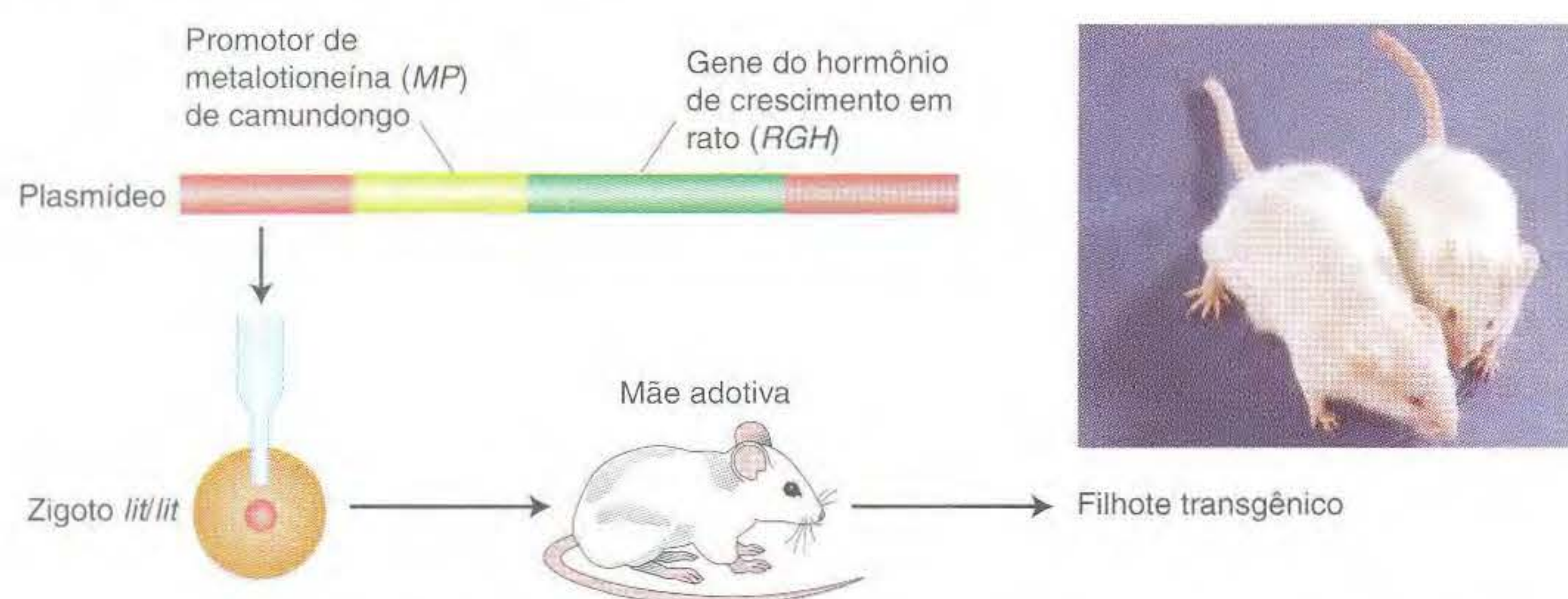
Injeção de transgenes Inserção aleatória e de homólogos no zigoto
Captação de transgenes Inserção aleatória e de homólogos pelas células-tronco

Nocautes gênicos direcionados:

Captação de transgene nulo por células-tronco Seleção de nocautes direcionados

Engenharia genética

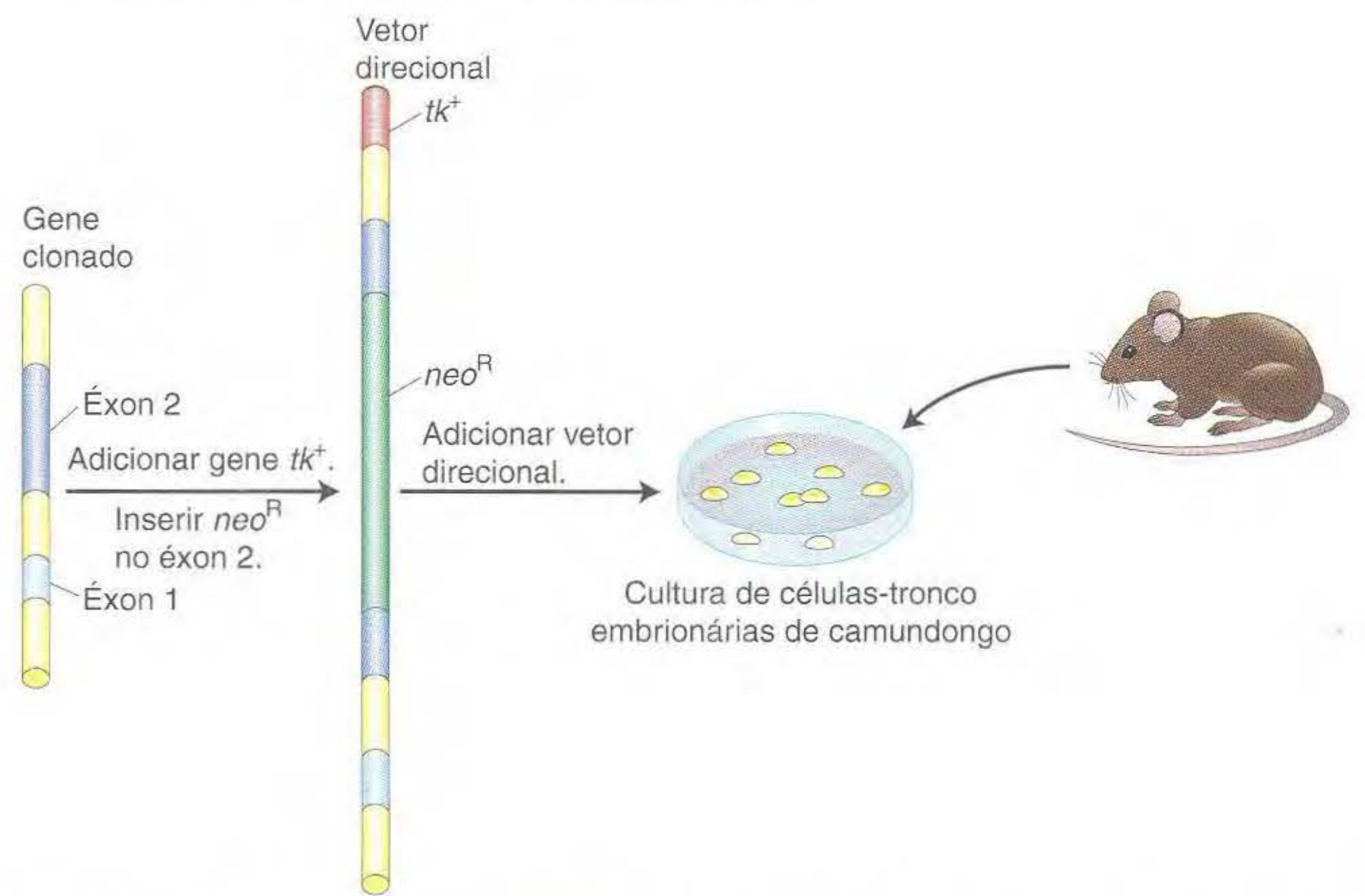
Transgênese. A criação de um camundongo transgênico é direta, mas requer uma manipulação cuidadosa de um zigoto (veja Capítulo 20). Primeiro, o DNA genômico de camundongo é clonado em *E. coli* com o uso de vetores bacterianos ou fagos. O DNA é então injetado em um zigoto, onde ele se integra em locais ectópicos (aleatórios) no genoma ou, menos comumente, no locus normal. A atividade da proteína transgênica pode ser monitorada fundindo-se o transgene com um gene repórter, tal como *GFP*, antes que o gene seja injetado. Com o uso de um método similar, as células somáticas de camundongo também podem ser modificadas por inserção de transgene: fragmentos específicos do DNA são inseridos em células somáticas individuais, e essas células são, por sua vez, inseridas nos embriões de camundongo.



Produção de um camundongo transgênico. O transgene, um gene de hormônio de crescimento de rato unido a um promotor de camundongo, é injetado em um zigoto de camundongo homozigoto para nanismo (*lit/lit*).

Nocautes direcionados. Os nocautes de genes específicos para dissecção genética podem ser feitos introduzindo-se um transgene contendo um alelo defeituoso e dois marcadores de resistência a drogas em uma célula-tronco embrionária tipo selvagem (veja Capítulo 20).

Produção de células ES com um nocaute gênico



Produção de um nocaute gênico. Um gene de resistência a drogas (neo^R) é inserido no transgene, servindo tanto como marcador como para perturbar o gene, produzindo um nocaute. (O gene tk é um segundo marcador.) A construção transgênica é então injetada nas células do embrião de camundongo.

Os marcadores são usados para selecionar as células transformantes específicas nas quais o alelo defeituoso substituiu o alelo tipo selvagem homólogo. As células transgênicas são então introduzidas em embriões de camundongo. Um método similar pode ser usado para substituir alelos tipo selvagem por um transgene funcional (terapia gênica).

Contribuições principais

No início da carreira do camundongo como organismo-modelo, os geneticistas usaram camundongos para elucidar os genes que controlam a cor e o padrão da pelagem, fornecendo um modelo para os mamíferos com pêlos, incluindo gatos, cães, cavalos e gado (veja Capítulo 6). Mais recentemente, os estudos de genética de camundongos forneceram uma gama de contribuições com significado direto para a saúde humana:

- Uma grande proporção de doenças genéticas humanas tem uma contraparte de camundongo, chamada de “modelo camundongo”, útil para estudos experimentais.
- Camundongos servem como modelos para os mecanismos de mutação de mamíferos.
- Os estudos dos mecanismos genéticos para o câncer são feitos em camundongos.
- Muitos carcinógenos potenciais são testados em camundongos.
- Camundongos foram modelos importantes para o estudo da genética do desenvolvimento de mamíferos (veja Capítulo 12). Por exemplo, eles fornecem um sistema-modelo para o estudo de genes que afetam a fenda labial e o palato fendido, um distúrbio desenvolvimental humano comum.
- As linhagens celulares que são fusões de híbridos de genomas de camundongo e humano têm um papel importante na localização de genes humanos em cromossomos humanos específicos. Há uma tendência para os cromossomos humanos serem perdidos de tais híbridos, e, assim, a perda de cromossomos específicos pode ser correlacionada à perda de alelos humanos específicos.

Outras áreas de contribuição

- Genética do comportamento
- Genética quantitativa
- Genes do sistema imunológico

Apêndice A

Nomenclatura Genética

Não existem regras universalmente aceitas para denominar os genes, alelos, produtos proteicos e fenótipos associados. No início, os geneticistas desenvolveram seus próprios símbolos para registrar seu trabalho. Depois, grupos de pessoas trabalhando com qualquer organismo reuniram-se e decidiram um conjunto de convenções que todos usariam. Como *Drosophila* foi um dos primeiros organismos a serem amplamente usados pelos geneticistas, a maioria dos atuais sistemas são variantes do sistema de *Drosophila*. Entretanto, houve uma considerável divergência. Alguns cientistas agora defendem uma padronização desse simbolismo, mas ela não foi obtida. Na verdade, a situação ficou mais complexa com o advento da tecnologia do DNA. Enquanto a maioria dos genes antes era denominada pelos fenótipos produzidos por mutações neles ocorridas, a nova tecnologia mostrou a natureza precisa dos produtos de muitos desses genes. Assim, parece mais apropriado chamá-los por sua função celular. Entretanto, os nomes antigos ainda estão na literatura, de modo que muitos genes têm dois conjuntos paralelos de nomenclatura.

Os exemplos seguintes não cobrem todos os organismos usados em Genética, mas a maioria dos sistemas segue um desses tipos.

<i>Drosophila melanogaster</i> (inseto)	
<i>ry</i>	Um gene que, quando mutado, causa olhos rosados
<i>ry</i> ⁵⁰²	Um alelo mutante recessivo específico que produz olhos rosados em homozigotos
<i>ry</i> ⁺	O alelo tipo selvagem de <i>rosy</i>
<i>ry</i>	O fenótipo mutante de <i>rosy</i>
<i>ry</i> ⁺	O fenótipo tipo selvagem (olhos vermelhos)
<i>RY</i>	O produto proteico do gene <i>rosy</i>
<i>XDH</i>	Xantina desidrogenase, uma descrição alternativa do produto proteico do gene <i>rosy</i> , chamado pela enzima que codifica
<i>D</i>	<i>Dichaete</i> ; um gene que, quando mutado, causa a perda de manutenção de algumas cerdas e asas lateralmente nos heterozigotos, causando letalidade nos homozigotos
<i>D</i> ³	Um alelo mutante específico do gene <i>Dichaete</i>
<i>D</i> ⁺	O alelo tipo selvagem de <i>Dichaete</i>
<i>D</i>	O fenótipo mutante <i>Dichaete</i>
<i>D</i> ⁺	O fenótipo tipo selvagem
<i>D</i>	(Dependendo do contexto) o produto proteico do gene <i>Dichaete</i> (uma proteína de ligação ao DNA)

<i>Neurospora crassa</i> (fungo)	
<i>arg</i>	Um gene que, quando mutado, causa necessidade de arginina
<i>arg-1</i>	Um gene <i>arg</i> específico
<i>arg-1</i>	Um alelo mutante inespecífico do gene <i>arg</i>
<i>arg-1</i> (1)	Um alelo mutante específico do gene <i>arg-1</i>
<i>arg-1</i> ⁺	O alelo tipo selvagem
<i>arg-1</i>	O produto proteico do gene <i>arg-1</i> ⁺

<i>Arg</i> ⁺	Uma linhagem que não requer arginina
<i>Arg</i> ⁻	Uma linhagem que requer arginina
<i>Saccharomyces cerevisiae</i> (fungo)	
<i>ARG</i>	Um gene que, quando mutado, causa necessidade de arginina
<i>ARG1</i>	Um gene <i>ARG</i> específico
<i>arg1</i>	Um alelo mutante inespecífico do gene <i>ARG</i>
<i>arg1-1</i>	Um alelo mutante específico do gene <i>ARG</i>
<i>ARG1</i> ⁺	O alelo tipo selvagem
<i>ARG1p</i>	O produto proteico do gene <i>ARG1</i> ⁺
<i>Arg</i> ⁺	Uma linhagem que não requer arginina
<i>Arg</i> ⁻	Uma linhagem que requer arginina
<i>Homo sapiens</i> (mamífero)	
<i>ACH</i>	Um gene que, quando mutado, causa acondroplasia
<i>ACH</i> ¹	Um alelo mutante (dominância não especificada)
<i>ACH</i>	Produto proteico do gene <i>ACH</i> ; natureza desconhecida
<i>FGFR3</i>	Nome recente do gene para acondroplasia
<i>FGFR3</i> ¹ ou <i>FGFR3</i> *1	
ou <i>FGFR3</i> <1>	Alelo mutante de <i>FGFR3</i> (inespecífico dominante)
Proteína <i>FGFR3</i>	Receptor 3 de fator de crescimento de fibroblasto
<i>Mus musculus</i> (mamífero)	
<i>Tyrc</i>	Um gene para tirosinase
⁺ <i>Tyrc</i>	O alelo tipo selvagem desse gene
<i>Tyrc</i> ^{ch} ou <i>Tyrc-ch</i>	Um alelo mutante causando a cor chinchilla
<i>Tyrc</i>	O produto proteico desse gene
<i>+TYRC</i>	O fenótipo tipo selvagem
<i>TYRCch</i>	O fenótipo chinchilla
<i>Escherichia coli</i> (bactéria)	
<i>lacZ</i>	Um gene para uso de lactose
<i>lacZ</i> ⁺	O alelo tipo selvagem
<i>lacZ1</i>	Um alelo mutante
<i>LacZ</i>	O produto proteico desse gene
<i>Lac</i> ⁺	Uma linhagem capaz de usar lactose (fenótipo)
<i>Lac</i> ⁻	Uma linhagem incapaz de usar lactose (fenótipo)
<i>Arabidopsis thaliana</i> (planta)	
<i>YGR</i>	Um gene que, quando mutante, produz folhas verde-amareladas
<i>YGR1</i>	Um gene <i>YGR</i> específico
<i>YGR1</i>	O alelo tipo selvagem
<i>ygr1-1</i>	Um alelo mutante recessivo específico de <i>YGR1</i>
<i>ygr1-2D</i>	Um alelo dominante específico (D) mutante de <i>YGR1</i>
<i>YGR1</i>	O produto proteico de <i>YGR1</i>
<i>Ygr</i> ⁻	Fenótipo verde-amarelado
<i>Ygr</i> ⁺	Fenótipo tipo selvagem

Apêndice B

Recursos de Bioinformática para Genética e Genômica

“Você certamente encontrará algo, se procurar, mas nem sempre é o que você procura” – *The Hobbit*, J.R.R. Tolkien

O campo da Bioinformática engloba o uso de computação para esclarecer dados complexos. Os dados de genética e de genômica são tão diversos que se tornou um considerável desafio identificar o(s) site(s) principal(is) para um tipo específico de informação. Além disso, o panorama de *software* acessível da Web para análise dessa informação está mudando constantemente à medida que são desenvolvidas novas e mais poderosas ferramentas. Este apêndice pretende dar *alguns* valiosos pontos de início para explorar o universo em rápida expansão de recursos *online* para genética e genômica.

1. Encontrar Web Sites de Genética e Genômica

Aqui estão listados vários recursos centrais que contêm grandes listas de Web sites relevantes:

- O periódico científico chamado *Nucleic Acids Research (NAR)* publica um número especial a cada janeiro listando uma ampla variedade de recursos de bancos de dados *online* em <http://nar.oupjournals.org/>
- O Virtual Library tem subdivisões de Model Organisms and Genetics com ricos recursos da Internet em <http://ceolas.org/VL/mo/> e http://www.ornl.gov/TechResources/Human_Genome/organisms.html
- O National Human Genome Research Institute (NHGRI) mantém uma lista de Web sites de genoma em <http://www.nhgri.nih.gov/10000375/>
- O Department of Energy (DOE) mantém um site do Human Genome Project no Oak Ridge National Laboratory em <http://public.ornl.gov/hgmis/>
- O SwissProt mantém Amos' WWW links em <http://www.expasy.ch/alinks.html>

2. Bancos de Dados Gerais

Bancos de Dados de Sequências de Proteínas e Ácidos Nucleicos Por acordo internacional, três grupos abrigam as sequências primárias de DNA e mRNA de todas as espécies: o National Center for Biotechnology Information (NCBI) abriga o GenBank; o European Bioinformatics Institute (EBI) abriga o European Molecular Biology Laboratory (EMBL) Data Library; e o National Institute of Genetics in Japan abriga o DNA DataBase do Japão (DDBJ).

Os registros primários de sequência de DNA, chamados acessos, são apresentados por grupos de pesquisa individuais. Além de dar acesso a esses registros de sequências de DNA, esses sites dão muitos outros conjuntos de dados. Por exemplo, o NCBI também abriga o RefSeq, um resumo de informações sobre as sequências de DNA de genomas totalmente sequenciados e os produtos gênicos que são codificados por essas sequências.

Muitas outras características importantes podem ser encontradas nos sites NCBI, EBI e DDBJ. *Home pages* e alguns outros sites importantes da Web são

- **NCBI** <http://www.ncbi.nlm.nih.gov/>
- **NCBI-Genomes** <http://www.ncbi.nlm.nih.gov/Genomes/index.html>
- **NCBI-RefSeq** <http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html>
- **The UCSC Genoma Bioinformatics Site** <http://genome.ucsc.edu/>

Esse ótimo site contém referências sobre sequências e montagens de rascunhos para uma grande coleção de genomas e muitas ferramentas para explorar esses genomas. O Genome Browser percorre os cromossomos, mostrando o trabalho de anotadores em todo o mundo. O Gene Sorter mostra expressão, homologia e outras informações sobre grupos de genes que podem ser relacionados de vários modos. Blat mapeia rapidamente sequências no genoma. O Table Browser dá acesso a bancos de dados subjacentes.

- **EBI** <http://www.ncbi.nlm.nih.gov/>
- **DDBJ** <http://www.nig.ac.jp/>

A dura realidade é que, com tantas informações biológicas, a meta de tornar esses recursos *online* “transparentes” ao usuário não foi totalmente obtida. Assim, a exploração desses sites irá envolver sua familiarização com os conteúdos de cada site e explorar alguns dos modos pelos quais o site o ajuda a enfocar suas dúvidas de modo a ter a resposta certa a suas perguntas. Para um exemplo do poder desses sites, considere uma pesquisa de uma sequência de nucleotídeos no NCBI. Os bancos de dados tipicamente guardam informações em bins separados chamados de “campos”. Usando dúvidas que limitam a pesquisa ao campo apropriado, pode ser feita uma pergunta mais direta. Usando a opção “Limits”, uma frase pode ser usada para identificar ou localizar determinada espécie, tipo de sequência (genômica ou mRNA), símbolo gênico ou qualquer outro campo de dados. As perguntas geralmente apóiam a habilidade em unir múltiplas dúvidas. Por exemplo: procure todos os registros de sequências de DNA da espécie *Caenorhabditis elegans* E que foram publicados após 1^a de janeiro de 2000. Usando a opção “History”, os resultados de muitas perguntas podem ser unidos, de modo que apenas os que forem comuns a várias perguntas serão recuperados. Com o uso apropriado das opções “*query*” em um site, muitos resultados falso-positivos podem ser computacionalmente eliminados sem descartar nenhum ponto relevante.

Como as previsões da sequência de proteínas são uma parte natural da análise de sequências de DNA e RNA, esses mesmos sites dão acesso a uma variedade de bancos de dados de proteínas. Um banco de dados importante é o SwissProt/TrEMBL. As sequências TrEMBL são automaticamente previstas de sequências de DNA e/ou mRNA. As sequências SwissProt são curadas, significando que experientes cientistas revisaram os resultados da análise computacional e tomaram decisões precisas sobre quais resultados aceitar ou rejeitar. Além dos registros primários de sequências de proteínas, o SwissProt também oferece bancos de dados de domínios de proteínas (fileiras de sequências de aminoácidos que são características de proteínas de um tipo particular). A *home page* do SwissProt é <http://www.ebi.ac.uk/swissprot/>.

Bancos de Dados de Domínios de Proteínas As unidades funcionais dentro das proteínas são tidas com sendo regiões de dobramento chamadas de domínios. A previsão dos domínios dentro de proteínas recém-descobertas é um modo de supor sua função. Vários bancos de dados de domínios de proteínas surgiram prevendo domínios que diferem de alguns modos. Alguns dos domínios de bancos de dados individuais são Pfam, PROSITE, PRINTS, SMART, ProDom, TIGRFAMs, BLOCKS e CDD. O InterPro permite consultar simultaneamente vários domínios de bancos de dados e apresentar resultados combinados. Web sites para alguns domínios de bancos de dados são

- **InterPro** <http://www.ebi.ac.uk/interpro/>
- **Pfam** <http://www.sanger.ac.uk/Software/Pfam/index.shtml>
- **PROSITE** <http://www.expasy.ch/prosite/>
- **PRINTS** <http://www.bioinf.man.ac.uk/dbbrowser/PRINTS/>
- **SMART** <http://smart.embl-heidelberg.de/>
- **ProDom** <http://prodes.toulouse.inra.fr/prodom/doc/prodom.html>
- **TIGRFAMs** <http://www.tigr.org/TIGRFAMs/>
- **BLOCKS** <http://blocks.fhcrc.org/>
- **CDD** <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>

Bancos de Dados de Estruturas de Proteínas A representação das estruturas tridimensionais das proteínas tornou-se um aspecto importante da análise molecular global. Os bancos de dados de estruturas tridimensionais estão disponíveis nos sites principais de sequências de DNA/proteína e bancos independentes de estrutura de proteínas, notadamente o Protein DataBase (PDB). O NCBI tem uma aplicação chamada Cn3D que ajuda na consulta de dados PDB.

- **PDB** <http://www.rcsb.org/pdb/>
- **Cn3D** <http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>

3. Bancos de Dados Especializados

Bancos de Dados Genéticos Específicos de Organismos Para reunir algumas classes de informação genética e genômica, principalmente informa-

ções fenotípicas, é necessário um conhecimento especial de determinada espécie. Assim, MOD (banco de dados de organismos-modelo) surgiram para fazer esse papel para os principais sistemas genéticos. Eles incluem bancos de dados de *Saccharomyces cerevisiae* (SGD), *Caenorhabditis elegans* (WormBase), *Drosophila melanogaster* (FlyBase), o peixe-zebra *Danio rerio* (ZFIN), o camundongo *Mus musculus* (MGI), o rato *Rattus norvegicus* (RGD), *Zea mays* (MaizeGDB) e *Arabidopsis thaliana* (TAIR). As *home pages* para esses MOD podem ser encontradas em

- **SGD** <http://genome-www.stanford.edu/Saccharomyces/>
- **WormBase** <http://www.wormbase.org/>
- **FlyBase** <http://flybase.org/>
- **ZFIN** <http://zfin.org/>
- **MGI** <http://www.informatics.jax.org/>
- **RGD** <http://rgd.mcw.edu/>
- **MaizeGDB** <http://www.maizegdb.org>
- **TAIR** <http://www.arabidopsis.org/>

Genética Humana e Bancos de Dados de Genômica Devido à importância da genética humana na clínica e na pesquisa básica, surgiram vários bancos de dados de genética humana. Entre eles está um banco de dados de doenças genéticas humanas chamado Online Mendelian Inheritance in Man (OMIM), um banco de dados com descrições resumidas de genes humanos chamado GeneCards, uma compilação de todas as mutações conhecidas em genes humanos chamada Human Gene Mutation Database (HGMD), um banco de dados do mapa atual de seqüências do genoma humano chamado de Golden Path e alguns links a bancos de dados de doenças genéticas humanas:

- **OMIM** <http://www3.ncbi.nlm.nih.gov/Omim/>
- **GeneCards** <http://mach1.nci.nih.gov/cards/index.html>
- **HGMD** <http://www.hgmd.org>
- **Golden Path** <http://genome.ucsc.edu/goldenPath/hgTracks.html>
- **Online Genetic Support Groups** <http://www.mostgene.org/support/index.html>
- **Genetic Disease Information** <http://www.geneticalliance.org/diseaseinfo/search.html>

Bancos de Dados de Projetos de Genoma Os projetos individuais de genoma também têm Web sites, onde apresentam seus resultados, em geral incluindo informações que não aparecem em nenhum outro Web site do mundo. Os maiores centros de genoma com fundos públicos incluem

- **Whitehead Institute/MIT Center for Genome Research** <http://www-genome.wi.mit.edu/>
- **Washington University School of Medicine Genome Sequencing Center** <http://genome.wustl.edu/>
- **Baylor College of Medicine Human Genome Sequencing Center** <http://www.hgsc.bcm.tmc.edu/>
- **Sanger Institute** <http://www.sanger.ac.uk/>
- **DOE Joint Genomics Institute** <http://www.jgi.doe.gov/>

4. Correlações de Genes Dentro de e Entre Bancos de Dados

Os produtos gênicos podem estar relacionados em virtude de compartilhar uma origem evolutiva comum, compartilhar uma função comum ou participar da mesma via.

BLAST: Identification of Sequence Similarities As evidências de uma origem evolutiva comum vieram da identificação de similaridades de seqüências entre duas ou mais seqüências. Uma das ferramentas mais importantes para identificar tais similaridades no BLAST (Basic Local Alignment Search Tool), que foi desenvolvido pelo NCBI. BLAST é de fato um conjunto de programas relacionados e bancos de dados nos quais as correspondências locais entre longos trechos de seqüências podem ser identificadas e classificadas. Uma pesquisa de um DNA similar ou seqüência de proteínas pelo BLAST é uma das primeiras coisas que um pesquisador faz com um gene recentemente seqüenciado. Bancos de dados de seqüências diferentes podem ser acessados e organizados por tipo de seqüência (genoma de referência, atualizações recentes, não-redundância, EST etc.), e uma espécie particular ou grupo taxonômico pode ser especificado. Uma rotina do BLAST correlaciona uma seqüência de nucleotídeos da pergunta traduzida em seis matrizes com uma seqüência de proteínas do banco de dados. Outra correlaciona uma seqüência indagada de proteína com a tradução de seis matrizes de uma seqüência de nucleotídeos do banco de dados. Outras rotinas de BLAST são adaptadas para identificar padrões de correspondência de seqüências curtas ou alinhamentos pareados, para triar segmentos de DNA do tamanho do genoma, e assim por diante, e podem ser acessados pelo mesma página do nível de cima:

- **NCBI-BLAST** <http://www.ncbi.nlm.nih.gov/BLAST/>

Bancos de Dados de Ontologia de Função Outro enfoque das correlações entre os produtos gênicos é a atribuição desses produtos a papéis funcionais baseados em evidência experimental ou previsão. Ter um modo comum de descrever esses papéis, independentemente do sistema experimental, é de grande importância. Um grupo de cientistas de bancos de dados diferentes estão trabalhando juntos para desenvolver um conjunto comum de termos hierarquicamente arrumados, uma ontologia, para *função* (evento bioquímico), *processo* (o evento celular para o qual a proteína contribui) e a *localização subcelular* (onde um produto está situado em uma célula) como modo de descrever as atividades de um produto gênico. Essa ontologia particular é chamada de Ontologia Gênica (GO), e muitos bancos de dados diferentes de produtos gênicos hoje incorporam termos GO. Uma descrição completa pode ser encontrada em

- <http://www.geneontology.org/>

Bancos de Dados de Vias Um outro modo de correlacionar produtos uns com os outros é atribuindo-lhes etapas nas vias bioquímicas ou celulares. Os diagramas de vias podem ser usados como modos organizados de apresentar correlações destes produtos uns com os outros. Algumas das tentativas mais avançadas em produzir tais bancos de dados de vias incluem a Kyoto Encyclopedia of Genes and Genomes (KEGG), Signal Transduction Database (TRANSPATH) e What Is There – Interactive Metabolic Database (WIT):

- **KEGG** <http://www.genome.ad.jp/kegg/>
- **TRANSPATH** <http://transpath.gbf.de/>
- **WIT** <http://wit.mcs.anl.gov/WIT2/>

Glossário

A *Veja adenina; adenosina.*

aberração cromossômica Qualquer tipo de mudança na estrutura ou número de cromossomos.

abrigo seguro Um sítio no genoma onde a inserção de um elemento de transposição improvavelmente causa uma mutação, evitando assim o dano ao hospedeiro.

acentuassomo A montagem macromolecular responsável pela interação de elementos acentuadores e as regiões promotoras dos genes.

ácido desoxirribonucleico *Veja DNA.*

ácido ribonucleico *Veja RNA.*

acridina laranja Um mutágeno que tende a produzir mutações de mudança de matriz de leitura.

adaptabilidade darwiniana A probabilidade relativa de sobrevivência e reprodução para um genótipo.

adaptabilidade dependente de frequência Diferenças de adaptabilidade cuja intensidade muda com a frequência relativa dos genótipos na população.

adaptabilidade independente de frequência Adaptabilidade que não depende de interações com outros membros da mesma espécie.

adaptabilidade média A média da adaptabilidade de todos os membros individuais de uma população.

adaptação No sentido evolutivo, alguma característica herdável do fenótipo de um indivíduo que melhora suas chances de sobrevivência e reprodução no ambiente existente.

adenina (A) Uma base purina que faz par com timina na dupla hélice do DNA.

adenina metilase Uma enzima que metila bases adenina antes da replicação, distinguindo assim os filamentos antigos dos filamentos complementares recém-sintetizados.

adenosina (A) Um nucleosídeo contendo adenina como sua base.

adenosina monofosfato cíclica (cAMP) Uma molécula contendo uma ligação diéster entre os átomos de carbono 3' e 5' da ribose do nucleotídeo. Esse nucleotídeo modificado não pode ser incorporado ao DNA nem ao RNA. Ele tem um papel principal como um sinal intracelular na regulação de vários processos.

adição *Veja mutação indel.*

ADP Adenosina difosfato.

AFB₁ *Veja aflatoxina B₁.*

aflatoxina B₁ (AFB₁) Um mutágeno que se liga a guanina, levando à clivagem da ligação entre guanina e desoxirribose, produzindo um sítio apurínico.

agente alquilante Um agente químico que pode adicionar grupos alquila (p. ex., grupos etil ou metil) a outra molécula; muitos mutágenos atuam por alquilação.

agente intercalar Um mutágeno que pode inserir-se entre as bases empilhadas no centro da dupla hélice do DNA, causando uma elevada taxa de *mutações indel*.

alça de deleção A alça formada na meiose pelo pareamento de um cromossomo normal e um cromossomo contendo uma deleção.

alça de deslocamento (alça D) Uma estrutura formada durante a recombinação homóloga quando a ponta 3' do filamento invasor desloca uma das cromátides não-irmãs não-danificadas. A cromátide deslocada forma a alça D.

alça de inversão Uma alça formada pelo pareamento meiótico de homólogos em uma inversão heterozigota.

alça em grampo Uma alça unifilamentar no RNA ou DNA formada por pontes de hidrogênio complementares dentro de um único filamento de RNA ou DNA.

alelismo múltiplo A existência de vários alelos conhecidos de um gene.

alelo Uma das formas diferentes de um gene que pode existir em um único locus.

alelo dominante Um alelo que expressa seu efeito fenotípico mesmo quando em heterozigose com um alelo recessivo; assim, se *A* é dominante em relação a *a*, então *A/A* e *A/a* têm o mesmo fenótipo.

alelo fixado Um alelo para o qual todos os membros da população em estudo são homozigotos, não existindo assim outros alelos para esse locus na população.

alelo letal Um alelo cuja expressão resulta na morte do organismo individual que o expressa.

alelo mutante Um alelo diferente do alelo encontrado no tipo padrão ou selvagem.

alelo nulo Um alelo cujo efeito é a ausência seja de um produto gênico normal, em nível molecular, seja da função normal, em nível fenotípico.

alelo recessivo Um alelo cujo efeito fenotípico não é expresso em um heterozigoto.

alelo subletal Um alelo que causa a morte de alguma parte (mas não de todas) dos indivíduos que o expressam.

alelos múltiplos As várias formas de um gene, diferindo em sua sequência de DNA e/ou expressão.

alongamento O estágio da transcrição que se segue à iniciação e precede o término.

alopoliplóide *Veja anfidiplóide.*

Alu Um curto elemento de transposição que constitui mais de 10% do genoma humano. Os elementos *Alu* são retroelementos que não codificam proteínas e, como tais, são elementos não-autônomos.

ambiente Uma combinação de todas as condições externas a um organismo que podem afetar a expressão de seu genoma.

aminoácido Um peptídeo; o bloco estrutural básico das proteínas (ou polipeptídeos).

aminoacil-tRNA sintetase Uma enzima que liga um aminoácido a um tRNA antes de seu uso na tradução. Existem 20 diferentes aminoacil-tRNA, um para cada aminoácido.

amniocentese Uma técnica para testar o genótipo de um embrião ou feto no útero com um risco mínimo para a mãe ou para a criança.

amostra Um pequeno grupo de membros individuais ou observações que pretendem ser representativas de uma população maior da qual o grupo foi colhido.

AMP Adenosina monofosfato.

amplificação A produção de muitas cópias de DNA a partir de uma região master do DNA.

anáfase Um estágio intermediário da divisão nuclear durante o qual os cromossomos são levados para os pólos da célula.

análise da tétrade O uso de *tétrades* (definição 2) para estudar o comportamento dos cromossomos e genes na meiose.

análise de heredogramas Dedução de herança monogênica de fenótipos humanos pelo estudo da prole de reproduções em uma família, em geral cobrindo várias gerações.

análise de variância Uma metodologia estatística que atribui a proporção da variância em uma população a causas diferentes e suas interações.

análogo de base Uma substância cuja estrutura molecular imita a de uma base do DNA; devido à mimetização, o análogo pode agir como um mutágeno.

andar no cromossomo Um método para a dissecação de grandes segmentos de DNA, no qual um segmento clonado de DNA, geralmente eucariótico, é usado para triar clones de DNA recombinante do mesmo banco de genoma para outros clones contendo sequências vizinhas.

anfidiplóide Um alopoliplóide; um poliplóide formado da união de dois conjuntos separados de cromossomos e sua duplicação subsequente.

antecedente Refere-se a uma sequência de DNA ou RNA situada no lado 5' de um ponto de referência.

anticódon Uma trinca de nucleotídeos em uma molécula de tRNA que se alinha a determinado códon no mRNA sob a influência do ribossomo; o aminoácido levado pelo tRNA é inserido a uma cadeia crescente de proteína.

anticorpo Uma molécula de proteína (imunoglobulina), produzida pelo sistema imunológico, que reconhece determinada substância (antígeno) e liga-se a ela.

antígeno nuclear de proliferação celular (PCNA) Parte do replissomo, o PCNA é a versão eucariótica da proteína deslizante procariótica.

antiterminalizador Uma proteína que promove a continuação da transcrição impedindo o término da transcrição em sítios específicos do DNA.

apoptose As vias celulares responsáveis pela morte celular programada.

arcabouço (1) A estrutura central de um cromossomo à qual o solenóide de DNA é ligado em alças; composto amplamente de topoisomerase. (2) No projeto do genoma, um conjunto ordenado de contigs no qual pode haver espaços não-seqüenciados conectados por seqüências pareadas pelas pontas.

asco Em um fungo, um saco que circunda uma tétrade ou uma óctade de ascósporos.

ascosporo Um esporo sexual de algumas espécies de fungos nos quais são encontrados esporos em um saco chamado asco.

atenuação Um mecanismo regulador no qual o nível de transcrição de um óperon (tal como *trp*) é reduzido quando o produto final de uma via (p. ex., triptofano) é suficiente; a etapa regulada após o início da transcrição.

atenuador Uma região da seqüência de RNA que forma estruturas secundárias alternativas que controlam o nível de transcrição de óperons atenuados.

ativadora Uma proteína que, quando ligada a um elemento de DNA regulador de ação cis, tal como um operador ou um acentuador, ativa a transcrição de um promotor adjacente.

ATP (trifosfato de adenosina) A “molécula energética” das células, sintetizada principalmente nas mitocôndrias e cloroplastos; a energia da quebra do ATP ativa muitas reações celulares importantes.

autofecundação Fertilizar ovócitos com espermatozóides do mesmo indivíduo.

automontagem A habilidade de algumas estruturas biológicas multiméricas em serem montadas a partir de suas partes componentes pelo movimento das moléculas e formação de ligações químicas fracas entre as superfícies que têm formas complementares.

autopoliploide Um poliploide formado pela duplicação de um único genoma.

auto-radiografia Um processo no qual materiais radioativos são incorporados a estruturas celulares, que são então colocadas perto de um filme ou emulsão fotográfica, formando assim um padrão no filme (um auto-radiograma) correspondendo ao local do composto radioativo dentro da célula.

auto-radiograma Um padrão de pontos escuros em um filme fotográfico revelado ou emulsão na técnica de auto-radiografia.

autossomo Qualquer cromossomo que não é um cromossomo sexual.

auxotrófica Uma linhagem de microrganismos que só irá proliferar quando o meio for suplementado com uma substância específica não-necessária aos organismos tipo selvagem (*compare com* **prototrófica**).

BAC Cromossomo artificial bacteriano; um plasmídeo F construído para agir como vetor de clonagem que pode carregar grandes insertos.

bactéria lisogênica Uma bactéria capaz de lise espontânea devido, por exemplo, ao desacoplamento de um profago do cromossomo bacteriano.

bacteriófago (fago) Um vírus que infecta bactérias.

balanceador Um cromossomo com várias inversões, usado para reter combinações alélicas favoráveis no homólogo não-invertido.

balanço gênico A idéia de que um fenótipo normal requer uma proporção 1:1 de genes no genoma.

bandas Faixas transversas em cromossomos reveladas por vários corantes.

bandas cromossômicas Faixas transversas nos cromossomos de muitos organismos, reveladas por procedimentos especiais de coloração.

barreiras de isolamento Um elemento de DNA que impede o espalhamento da heterocromatina, servindo como um sítio de ligação para proteínas que mantêm modificações de cromatina eucromática, tais como acetilação de histona.

base *Véja* **base nucleotídica**.

base danificada oxidativamente Uma origem dos danos espontâneos ao DNA.

base nucleotídica A base de nitrogênio (um grupo purina ou pirimidina) que forma parte de um nucleotídeo.

bases complementares (pares de bases) Referem-se a pareamentos específicos entre adenina e timina e entre guanina e citosina.

bases nitrogenadas Tipos de moléculas que formam partes importantes de ácidos nucleicos, compostas de estruturas em anel contendo nitrogênio; pontes de hidrogênio entre bases ligadas a dois filamentos de uma dupla hélice de DNA.

biblioteca Uma coleção de clones de DNA obtidos de um DNA doador.

biblioteca de cDNA Uma biblioteca composta de cDNA, não necessariamente representando todos os mRNA.

biblioteca de expressão Uma biblioteca na qual o vetor leva sinais transcrpcionais para permitir que qualquer inserto clonado produza mRNA e, finalmente, uma proteína.

biblioteca genômica Uma biblioteca que engloba todo o genoma.

bioinformática Sistemas de informação computacional e métodos analíticos aplicados a problemas biológicos tais como análise genômica.

biologia de sistemas Uma tentativa de interpretar um genoma como um sistema holístico de interação.

bivalentes Dois cromossomos homólogos pareados na meiose.

bolha de transcrição O sítio no qual a dupla hélice é desenrolada de modo que a RNA polimerase possa usar um dos filamentos de DNA como molde para a síntese de RNA.

C *Véja* **citidina; citosina**.

caça *Véja* **experimento de pulso-caça**.

caça ao mutante O processo de coletar mutantes diferentes mostrando anomalias em certa estrutura ou em certa função, como preparação para a dissecação mutacional dessa função.

CAF-1 (fator 1 de montagem de cromatina) Uma proteína que se liga a histonas e as leva para a forquilha de replicação para sua incorporação em novos nucleossomos.

camada germinativa Uma das camadas primárias de células embrionárias que são formadas como resultado da gastrulação. As camadas germinativas primárias em animais superiores são o endoderma, mesoderma e ectoderma.

“caminhada” do primer (*primer walking*) O uso de um *primer* com base em uma área seqüenciada de um genoma para seqüenciar uma área flaqueadora não-seqüenciada.

cAMP *Véja* **adenosina monofosfato cíclica**.

câncer Uma classe de doença caracterizada pela proliferação rápida e descontrolada de células dentro de um tecido de um eucarionte com vários tecidos. Os cânceres são geralmente tidos como sendo doenças genéticas de células somáticas, surgindo por mutações seqüenciais que criam oncogenes e inativam genes supressores tumorais.

cap Uma estrutura especial, consistindo em uma 7-metilguanosina ligada ao transcrito por três grupos fosfato, que é adicionada no núcleo à ponta 5' do mRNA eucariótico. Cap protege um mRNA da degradação, e é necessário para a tradução do mRNA no citoplasma.

CAP *Véja* **proteína ativadora do catabolismo**.

característica (1) Um atributo dos membros individuais de uma espécie para os quais podem ser definidas várias diferenças herdáveis. (2) Mais ou menos sinônimo de fenótipo.

característica canalizada Uma característica que é constante no desenvolvimento a despeito de perturbações ambientais e genéticas.

característica familiar Uma característica comum aos membros de uma família.

carcinógeno Uma substância que causa câncer.

carga genética O conjunto total de alelos deletérios em um genótipo individual.

cariótipo O complemento cromossômico inteiro de um organismo individual ou célula, visto durante a metáfase mitótica.

casamento preferencial positivo Uma situação na qual fenótipos similares reproduzem-se mais comumente que o esperado pelo acaso.

cascata de sinal de transdução Uma série de eventos seqüenciais – tais como as fosforilações de proteínas, que passam um sinal recebido por um receptor transmembranar através de uma série de moléculas intermediárias até as moléculas reguladoras finais, tais como fatores de transcrição – que são modificados em resposta ao sinal.

cauda de histona O final de uma proteína histona que se projeta do cerne do nucleossomo e está sujeita a modificação pós-traducional. *Veja também código de histona.*

cauda poli(A) Uma sequência de nucleotídeos adenina ligada ao mRNA após a transcrição.

cDNA *Veja* **DNA complementar.**

célula com alta frequência de recombinação (Hfr) Em *E. coli*, uma célula tendo seu fator de fertilidade integrado ao cromossomo bacteriano; uma célula doadora (masculina).

célula eucariótica Uma célula contendo um núcleo.

célula F⁺ Em *E. coli*, uma célula tendo um fator de fertilidade livre; uma célula masculina.

célula F⁻ Em *E. coli*, uma célula sem fator de fertilidade; uma célula feminina.

célula procariótica Uma célula sem membrana nuclear e, assim, sem núcleo separado.

célula somática Uma célula que não é destinada a tornar-se um gameta; uma “célula do corpo” cujos genes não serão passados para as gerações futuras.

células ES *Veja* **células-tronco embrionárias.**

células-filhas Duas células idênticas formadas pela divisão assexual de uma célula.

célula-tronco Uma célula que se divide, em geral assimetricamente, para dar origem a duas células diferentes da prole. Uma célula da prole é um blastócito como a célula parental e a outra é uma célula que entra em uma via de diferenciação. Desse modo, uma célula de propagação contínua pode manter-se e gerar células diferenciadas.

células-tronco embrionárias (ES) Linhagens celulares cultivadas que são estabelecidas de embriões bem iniciais e são essencialmente totipotentes. Isto é, essas células podem ser implantadas em um embrião hospedeiro e povoar muitos dos ou todos os tecidos do animal em desenvolvimento, incluindo, sobretudo, a linhagem germinativa.

As manipulações dessas células ES são amplamente usadas em genética de camundongos para produzir nocautes gênicos direcionados.

centimorgan (cM) *Veja* **unidade de mapa.**

centro decodificador A região na subunidade ribossômica menor onde é tomada a decisão se um aminoacil-tRNA pode se ligar ao sítio A. Essa decisão é baseada na complementariedade entre o anticódon do tRNA e o códon do mRNA.

centro peptidiltransferase O sítio na subunidade ribossômica maior no qual é catalisada ligação de dois aminoácidos.

centrômero Uma região especializada do DNA de cada cromossomo eucariótico que atua como um sítio para a ligação das proteínas do cinetócoro.

cerne da enzima Como usado no Capítulo 8, todas as subunidades da RNA polimerase, exceto o fator σ .

chaperona Uma proteína que ajuda a dobrar corretamente as proteínas recém-sintetizadas. As chaperonas são conservadas em todos os organismos, desde bactérias até plantas e humanos.

ciclo celular Um conjunto de eventos que ocorrem nas divisões de células mitóticas. O ciclo celular oscila entre mitose (fase M) e intérfase. A intérfase pode ser subdividida em ordem em fase G₁, S e G₂. A síntese de DNA ocorre na fase S. A duração do ciclo celular é regulada por uma opção especial em G₁, na qual as células em G₁ podem entrar em uma fase de repouso chamada G₀.

ciclo lítico O ciclo de vida do bacteriófago que leva à lise da célula hospedeira.

cinase Uma enzima que adiciona um ou mais grupos fosfato a seus substratos. Se os substratos são proteínas específicas, a enzima é chamada uma cinase proteica.

cinase proteica Uma enzima que fosforila aminoácidos específicos em proteínas-alvo específicas.

cinetócoro Um complexo de proteínas ao qual se ligam as fibras do fuso nuclear.

citidina (C) Um nucleosídeo contendo citosina como sua base.

citoesqueleto O sistema de cabos de proteína e proteínas associadas que, juntos, formam a arquitetura de uma célula eucariótica.

citogenética O enfoque citológico da genética, consistindo principalmente em estudos microscópicos dos cromossomos.

citohet Uma célula contendo dois tipos geneticamente distintos de uma organela específica.

citoplasma O material entre a membrana nuclear e as membranas celulares; inclui o líquido (citossol), organelas e várias membranas.

citossina (C) Uma base pirimidínica que faz par com guanina.

citotipo M O nome dado a uma linhagem de *Drosophila* sem os elementos de transposição P.

citotipo P O nome dado a uma linhagem de *Drosophila* com elementos de transposição P.

clonagem (1) Na pesquisa do DNA recombinante, o processo de criar e amplificar segmentos específicos de DNA. (2) A produção de organismos geneticamente idênticos a partir de células somáticas de um organismo individual.

clonagem posicional A identificação de sequências de DNA codificando um gene de interesse baseado no conhecimento de sua localização genética ou de mapa citogenético.

clone (1) Um grupo de células geneticamente idênticas ou organismos individuais derivados por divisão assexual de um ancestral comum. (2) (*coloquial*) Um organismo individual formado por algum processo sexual de modo que seja geneticamente idêntico a seu “genitor”. (3) *Veja* **clone de DNA.**

clone contig *Veja* **contig.**

clone de DNA Um trecho do DNA que foi inserido em uma molécula vetor, tal como um plasmídeo ou um cromossomo de fago, e então replicado para formar muitas cópias.

cM (centimorgan) *Veja* **unidade de mapa.**

coativador Uma classe especial de complexo regulador eucariótico que serve como uma ponte para juntar proteínas regulatórias e RNA polimerase II.

código de histona Refere-se ao padrão de modificação (p. ex., acetilação, metilação, fosforilação) das caudas de histonas que podem levar a informação necessária para a montagem correta da cromatina.

código genético Um conjunto de correspondências entre trincas de pares de nucleotídeos no DNA e aminoácidos na proteína.

código redundante Um código genético no qual alguns aminoácidos podem ser codificados por mais de um códon cada.

co-dominância Uma situação na qual um heterozigoto apresenta os efeitos fenotípicos de ambos os alelos igualmente.

códon Um trecho do RNA (do tamanho de três pares de nucleotídeos) que codifica um único aminoácido.

códon âmbar O códon AUG, um códon sem sentido.

códon ocre O códon UAA, um códon sem sentido.

códon opala O códon UGA, um códon sem sentido.

códon sem sentido Um códon para o qual não existe molécula de tRNA formal; a presença de um códon sem sentido causa o término da tradução (final da cadeia polipeptídica). Os três códons sem sentido são chamados âmbar, ocre e opala.

coeficiente de coincidência A proporção do número observado de recombinantes duplos em relação ao número esperado.

coeficiente de correlação Uma medida estatística da extensão na qual as variações em uma variável estão relacionadas a variações em outra.

coeficiente de endogamia A probabilidade de homozigose que resulta porque o zigoto obtém cópias do *mesmo* gene ancestral.

coeficiente de regressão A inclinação da linha reta que mais proxima-mente se relaciona com duas variantes correlatas.

cointegrado O produto de fusão de dois elementos circulares transponíveis para formar um único círculo maior na transposição replicativa.

colinearidade A correspondência entre a localização de um sítio mutante dentro de um gene e a localização de uma substituição de aminoácido dentro do polipeptídeo traduzido desse gene.

colônia Um clone visível de células.

compensação de dose O processo em organismos que usam um mecanismo cromossômico de determinação do sexo (tal como um XX *versus* XY) que permite que genes estruturais padrão no cromossomo sexual sejam

expressos nos mesmos níveis nas mulheres e homens, independentemente do número de cromossomos sexuais. Em mamíferos, a compensação de dose opera mantendo apenas um cromossomo X ativo em cada célula; em *Drosophila*, ele opera hiperativando o cromossomo X do macho.

complementação A produção de um fenótipo tipo selvagem quando duas mutações diferentes são combinadas em um diplóide ou um heterocarion.

complementação funcional O uso de um fragmento clonado de DNA tipo selvagem para transformar um mutante em tipo selvagem; usada na identificação de um clone contendo um gene específico.

complexo de gene homeótico Um dos grupos fortemente ligados de genes que controlam a determinação do segmento ântero-posterior nos animais segmentados. Todos esses genes codificam fatores de transcrição de homeodomínio.

complexo de pré-iniciação (PIC) Um complexo proteico eucariótico muito grande constituído de uma RNA polimerase II e seis fatores gerais de transcrição (GTF), cada um dos quais é um complexo multiproteico.

complexo de reconhecimento da origem (ORC) Um complexo proteico que se liga às origens de replicação das leveduras.

complexo gênico Um grupo de genes adjacentes, funcional e estruturalmente relacionados, que surgem tipicamente por duplicação gênica no curso da evolução.

complexo mediador Um complexo proteico que atua como adaptador interagindo com fatores de transcrição ligados a sítios reguladores e com fatores gerais de iniciação para a transcrição mediada pela RNA polimerase II.

complexo sinaptonêmico Uma estrutura complexa que une homólogos durante a prófase da meiose.

complexo ternário Um complexo contendo um tRNA aminoacilado (tRNA mais aminoácido) e o fator de alongamento Tu (EF-Tu). O complexo ternário liga-se ao sítio A de um ribossomo.

composto ICR Um de uma série de compostos de quinacrina mostarda sintetizados no Institute for Cancer Research (Fox Chase, Pennsylvania) que atuam como *agentes intercalantes*.

condição permissiva Uma condição ambiental sob a qual um mutante condicional apresenta o fenótipo tipo selvagem.

condição restritiva Condição ambiental sob a qual um mutante condicional mostra um fenótipo mutante.

conformação cis Observada em um heterozigoto tendo dois sítios mutantes dentro de um gene ou dentro de um grupo de genes; o arranjo A_1A_2/a_1a_2 .

conformação trans Em um heterozigoto com dois sítios mutantes dentro de um grupo de genes, o arranjo $a_1 +/+ a_2$.

conjugação A união de duas células bacterianas durante a qual o material cromossômico é transferido da célula doadora para a receptora.

conjunto de cromossomos O grupo de cromossomos diferentes que levam a informação genética básica de determinada espécie.

constitutiva Refere-se a uma função biológica que está sempre em um estado, independentemente das condições ambientais. (1) Com relação à expressão gênica, refere-se a um gene que é sempre expresso. (2) Em relação à estrutura do cromossomo, refere-se a partes do cromossomo que são sempre heterocromáticas.

contig Um conjunto de clones ordenados superpostos que constituem uma região cromossômica de um genoma.

controle negativo Regulação mediada por fatores que bloqueiam ou desligam a transcrição.

controle positivo Regulação mediada por uma proteína que é necessária para a ativação de uma unidade de transcrição.

conversão gênica Um processo meiótico de mudança dirigida no qual um alelo dirige a conversão de um alelo para sua própria forma.

cooperatividade A necessidade de proteínas interagirem para funcionar. A cooperatividade de fatores de transcrição é uma característica da regulação gênica eucariótica.

co-repressor Uma proteína ou complexo proteico que atua em conjunto com uma proteína repressora para reprimir a transcrição, em geral modificando a estrutura da cromatina; geralmente não é uma entidade de ligação ao DNA em si.

corpúsculo de Barr Uma massa densamente corada que representa um cromossomo X inativado.

correção A produção (possivelmente por excisão e reparo) de um par de nucleotídeos apropriadamente pareados de uma sequência de DNA híbrido que contém um par ilegítimo.

correlação A tendência de uma variável em variar em proporção a outra variável, seja positiva ou negativamente.

correlação fenotípica Similaridade fenotípica entre grupos a despeito de essa similaridade ser o resultado de semelhança genética.

correlação genética Similaridade fenotípica entre grupos diferentes porque os grupos herdaram os mesmos genes.

co-segregação Herança paralela de dois genes devido à sua próxima ligação em um cromossomo.

cosmídeo Um vetor de clonagem que pode replicar-se autonomamente como um plasmídeo e ser embalado em um fago.

co-supressão Um fenômeno epigenético em que um transgene torna-se reversivelmente inativado juntamente com a cópia do gene no cromossomo.

co-transdução A transdução simultânea de dois genes marcadores bacterianos.

co-transdutantes Uma célula bacteriana transduzida simultaneamente por dois alelos doadores. Sua frequência é usada como uma medida de proximidade dos genes doadores no mapa cromossômico.

co-transformação A transformação simultânea de dois genes marcadores bacterianos.

covariança Uma medida estatística usada para computar o coeficiente de correlação entre duas variáveis; a covariança é a média de $(x - \bar{x})(y - \bar{y})$ de todos os pares de valores para as variáveis x e y , onde \bar{x} é a média dos valores x e \bar{y} é a média dos valores y .

cpDNA DNA de cloroplasto.

cristalografia de raios X Uma técnica para deduzir a estrutura molecular direcionando um feixe de raios X para um cristal do composto testado e medindo a dispersão dos raios.

cromátide Uma das duas réplicas lado a lado produzidas pela divisão cromossômica.

cromátides-irmãs O par justaposto de cromátides que surge da replicação de um cromossomo.

cromatina A substância dos cromossomos; agora conhecida como incluindo DNA e proteínas cromossômicas.

cromômero Uma pequena estrutura em forma de conta visível em um cromossomo durante a prófase da meiose e mitose.

cromossomo Um arranjo linear de ponta a ponta de genes e outro DNA, às vezes com proteínas associadas e RNA.

cromossomo acêntrico Um cromossomo sem centrômero.

cromossomo acrocêntrico Um cromossomo que tem o centrômero situado bem mais perto de uma ponta que da outra.

cromossomo artificial bacteriano Veja BAC.

cromossomo artificial de levedura (YAC) Um sistema vetor de clonagem em *Saccharomyces cerevisiae* empregando o centrômero de levedura e sequências de replicação.

cromossomo dicêntrico Um cromossomo com dois centrômeros.

cromossomo politênico Um cromossomo gigante em tecidos específicos de alguns insetos, produzido por um processo endomitótico no qual os múltiplos conjuntos de DNA permanecem ligados em um número haplóide de cromossomos.

cromossomo satélite Um cromossomo que parece ser uma adição ao genoma normal.

cromossomo X Um de um par de cromossomos sexuais, distinto do cromossomo Y.

cromossomo Y Um cromossomo do par de cromossomos sexuais distinto do cromossomo X.

cromossomos homeólogos Cromossomos parcialmente homólogos, geralmente indicando alguma homologia ancestral original.

cromossomos homólogos Cromossomos que fazem pares entre si na meiose ou cromossomos em espécies diferentes que retiveram a maioria dos mesmos genes durante sua evolução de um ancestral comum.

cromossomos sexuais Um cromossomo cuja presença/ausência é correlacionada ao sexo do portador; um cromossomo que tem um papel na determinação do sexo.

crossing over A troca de partes cromossômicas correspondentes entre homólogos por quebra e reunião.

cruzamento A reprodução deliberada de dois tipos parentais de organismos em análise genética.

cruzamento diíbrido Um cruzamento entre dois indivíduos identicamente heterozigotos em dois loci; por exemplo, $A B/a b \times A B/a b$.

cruzamento monoíbrido Um cruzamento entre dois heterozigotos idênticos em um par de genes – por exemplo, $A/a \times A/a$.

cruzamentos recíprocos Um par de cruzamentos do tipo genótipo A (fêmea) \times genótipo B (macho) e B (fêmea) \times A (macho).

cruzamento-teste Um cruzamento de um organismo individual de genótipo desconhecido ou um heterozigoto (ou um heterozigoto múltiplo) com um *testador*.

cruzamento-teste de três pontos Um cruzamento-teste no qual um genitor tem três pares de genes heterozigotos.

CTD *Veja domínio da cauda carboxila.*

cultura Tecido ou células que se multiplicam por divisão assexual e crescem para experimentação.

CVS *Veja punção de vilosidades coriônicas.*

deleção A remoção de um segmento cromossômico em um conjunto de cromossomos.

deleção intragênica Uma deleção dentro de um gene.

deleção multigênica Uma deleção de vários genes adjacentes.

deriva *Veja deriva genética aleatória.*

deriva genética A mudança de frequência de um alelo em uma população resultante de diferenças casuais no número de prole de genótipos diferentes produzidos por membros individuais diferentes.

deriva genética aleatória Mudanças em frequências alélicas que ocorrem porque os genes que aparecem na prole não são uma amostra perfeitamente representativa dos genes parentais.

desaminação A remoção de grupos amino dos compostos. A desaminação de citosina produz uracil. Se essa desaminação não for reparada, pode ocorrer transição de citosina para timina.

descoberta de gene Um processo pelo qual os geneticistas encontram um conjunto de genes que afetam alguns processos biológicos de interesse por padrões de herança monogênica de seus alelos mutantes ou por análise genômica.

desenvolvimento O processo pelo qual uma única célula torna-se um organismo diferenciado.

desequilíbrio de ligação Desvio na frequência de haplótipos em uma população da frequência esperada se os alelos em loci diferentes estiverem associados aleatoriamente.

desnaturação A separação de dois filamentos de uma dupla hélice de DNA ou a grave perturbação da estrutura de qualquer molécula complexa sem quebra das principais ligações de suas cadeias.

desoxirribose O açúcar pentose no arcabouço do DNA.

destino celular O estado diferenciado final ao qual a célula se destina.

desvio padrão A raiz quadrada da variância.

determinante Uma molécula localizada espacialmente que faz com que as células adotem um destino particular ou conjunto de destinos relacionados.

diáde Um par de cromátides-irmãs unidas no centrômero, como na primeira divisão da meiose.

Dicer Um complexo proteico que reconhece longas moléculas de RNA bifilamentares e as cliva em siRNA bifilamentares. Dicer tem um papel central na interferência de RNA.

diferença de característica Formas alternativas do mesmo atributo dentro de uma espécie.

diferenciação As mudanças na forma e fisiologia celular associadas à produção dos tipos celulares finais de determinado órgão ou tecido.

diíbrido Um duplo heterozigoto tal como $A/a \cdot B/b$.

dímero de timina Um par de bases timina adjacentes quimicamente ligadas no DNA; os processos celulares que reparam essa lesão em geral cometem erros que criam mutações.

dimorfismo Um polimorfismo apenas com duas formas.

diplóide Uma célula que tem dois conjuntos cromossômicos, ou um organismo individual tendo dois conjuntos cromossômicos em cada uma de suas células.

diplóide parcial *Veja merozigoto.*

diplóide transiente O estágio do ciclo de vida de fungos (e algas) predominantemente haplóides no qual ocorre a meiose.

direcionamento Uma característica de alguns elementos de transposição que facilita sua inserção em regiões do genoma onde provavelmente eles não se inserem em um gene causando uma mutação.

disgenesia híbrida Uma síndrome de efeitos incluindo esterilidade, mutação, quebra cromossômica e recombinação masculina na prole híbrida de crossings entre algumas drosófilas de laboratório e isolados naturais.

dispersão A variação de valores de alguma variável em uma população ou amostra.

dissecção genética O uso de recombinação e mutação para juntar os vários componentes de determinada função biológica.

dissecção mutacional O estudo dos componentes de uma função biológica por um estudo de mutações que afetam essa função.

dissociação A separação de uma molécula de DNA em dois filamentos. Também chamada de *desnaturação*.

dissômico Um haplóide anormal que leva duas cópias de um cromossomo.

distribuição *Veja distribuição estatística.*

distribuição bimodal Uma distribuição estatística tendo duas modas.

distribuição de equilíbrio Uma disposição de frequências genotípicas ou fenotípicas em uma população que permanece constante com o tempo.

distribuição de Poisson Uma expressão matemática que dá a probabilidade de observar vários números de um evento particular em uma amostra quando a probabilidade média de um evento em qualquer tentativa é muito pequena.

distribuição estatística O arranjo de frequências de classes diferentes quantitativa ou qualitativamente em uma população.

distribuição independente *Veja Segunda lei de Mendel.*

distúrbio genético heterogêneo Um distúrbio causado por mutações em qualquer um dos vários genes que codificam determinado processo.

ditipo não-parental (NPD) Uma tétrade contendo dois genótipos diferentes, ambos os quais são recombinantes.

ditipo parental (PD) Um tipo de tétrade contendo dois genótipos diferentes, ambos os quais são parentais.

divisão celular Um processo pelo qual duas células são formadas a partir de uma.

DNA (ácido desoxirribonucleico) Uma cadeia de nucleotídeos ligados (tendo a desoxirribose como seu açúcar). Duas de tais cadeias em forma de dupla hélice são a substância fundamental da qual são compostos os genes.

DNA complementar (cDNA) DNA sintético transcrito de um RNA específico pela ação da enzima transcriptase reversa.

DNA de cloroplasto (cpDNA) O pequeno componente genômico encontrado nos cloroplastos das plantas, envolvido com a fotossíntese e outras funções que ocorrem dentro dessa organela.

DNA de cópia única Sequências de DNA presentes apenas em uma cópia por genoma haplóide.

DNA doador Qualquer DNA a ser usado na clonagem ou na transformação mediada pelo DNA.

DNA espaçador DNA encontrado entre os genes; sua função é desconhecida.

DNA exógeno DNA de outro organismo.

DNA glicosilase Uma enzima que remove bases alteradas, deixando sítios apurínicos.

DNA hemimetilado Sequência de DNA com um filamento metilado e um filamento não-metilado.

DNA ligase Uma enzima importante na replicação e reparo do DNA que sela o arcabouço de DNA catalisando a formação de ligações fosfodiéster.

DNA microssatélite Um tipo de DNA repetitivo que consiste em repetições muito curtas, tais como dinucleotídeos.

DNA mitocondrial (mtDNA) O subgrupo do genoma encontrado nas mitocôndrias, especializado em fornecer algumas das funções da organela.

DNA polimerase Uma enzima que catalisa a síntese de novos filamentos de DNA a partir de um molde de DNA; existem várias dessas enzimas.

DNA recombinante Uma nova sequência de DNA formada pela combinação de duas moléculas não-homólogas de DNA.

DNA redundante *Veja DNA repetitivo.*

DNA repetitivo DNA redundante; sequências de DNA que estão presentes em muitas cópias por conjunto cromossômico.

DNA satélite Qualquer tipo de DNA altamente repetitivo; antes definido como DNA formador de uma banda satélite após centrifugação em um gradiente de densidade de cloreto de cério.

DNA sucata (*junk*) Nome dado aos elementos de transposição pelos que acreditam que eles são DNA parasitas e não desempenham função útil para o hospedeiro.

DnaA Uma proteína de *E. coli*. A replicação do DNA de *E. coli* começa quando a proteína DnaA liga-se a uma sequência de 13 pb (chamada de DnaA box) na origem da replicação.

DnaB Uma proteína de *E. coli* que abre a dupla hélice na forquilha de replicação. O DnaB é conhecido como uma helicase de DNA.

doador Bactéria usada em estudos de transmissão unidirecional de DNA para outras células; os exemplos são Hfr em conjugação e a fonte do fago na transdução.

dominância completa Descreve um alelo que se expressa da mesma maneira tanto em uma cópia única (heterozigoto) como em uma cópia dupla (homozigoto).

dominância incompleta Uma situação na qual um heterozigoto apresenta um fenótipo quantitativamente (mas não exatamente) intermediário entre os fenótipos homozigotos correspondentes. (Exatamente intermediário significa sem dominância.)

dominância total *Veja dominância completa.*

dominante negativo Um alelo dominante que, em dose única (um heterozigoto), elimina a função do gene por um efeito de prejudicar a proteína.

domínio Região de uma proteína associada a determinada função. Algumas proteínas contêm mais de um domínio.

domínio da cauda carboxila (CTD) A cauda de proteína da subunidade β da RNA polimerase II; ela coordena o processamento dos pré-mRNA eucarióticos, incluindo capping, recomposição e término.

domínio de ativação Parte de um fator de transcrição necessária para a ativação da transcrição do gene-alvo; pode ligar-se a componentes da maquinaria transcrricional e/ou recrutar proteínas que modificam a estrutura da cromatina.

domínio de ligação ao DNA O sítio em uma proteína de ligação ao DNA que interage diretamente com sequências específicas de DNA.

dosagem gênica O número de cópias de determinado gene presente no genoma.

dupla hélice A estrutura do DNA primeiro proposta por James Watson e Francis Crick, com duas hélices entrelaçadas unidas por pontes de hidrogênio entre bases pareadas.

duplicação Mais de uma cópia de determinado segmento cromossômico em um conjunto cromossômico.

duplicação de sítio-alvo Uma curta sequência de DNA de repetição direta (tipicamente de 2 a 10 pb de tamanho) adjacente às pontas de um elemento de transposição que foi gerado durante a integração do elemento no cromossomo hospedeiro.

duplicação em tandem Segmentos cromossômicos adjacentes idênticos.

duplicação insercional Uma duplicação na qual a cópia extra não é adjacente à normal.

duplo crossing Dois crossings em uma região cromossômica em estudo.

efeito aditivo Uma diferença no fenótipo resultante da substituição de um alelo *a* por um alelo *A* na média de todos os membros individuais de uma população.

efeito de dosagem gênica (1) Proporcionalidade da expressão de alguma função biológica do número de cópias de um alelo presente na célula. (2) Uma mudança no fenótipo causada por um número anormal de alelos tipo selvagem (observado nas mutações cromossômicas).

efeito de posição Descreve a situação na qual a influência fenotípica de um gene é alterada por mudanças na posição do gene dentro do genoma.

efeito do fundador Uma diferença aleatória da população parental na frequência de um genótipo em uma nova colônia que resulta de um pequeno número de fundadores.

efeito materno A influência ambiental dos tecidos maternos no fenótipo da prole.

efeito sinérgico Uma característica das proteínas regulatórias eucarióticas para as quais a ativação transcrricional mediada pela interação de várias proteínas é maior do que a soma dos efeitos das proteínas avaliadas individualmente.

efetor alostérico Uma pequena molécula que se liga a um sítio alostérico.

EF-G *Veja fator de alongamento G.*

EF-Tu *Veja fator de alongamento Tu.*

elemento Ac *Veja elemento ativador.*

elemento acentuador Uma sequência regulatória de ação cis que pode elevar os níveis de transcrição de um promotor adjacente. Muitos acentuadores histoespecíficos podem determinar padrões espaciais de expressão gênica em eucariontes superiores. Os acentuadores podem agir nos promotores a dezenas de quilobases de DNA e podem estar a 5' ou 3' dos promotores que eles regulam.

elemento ativador (Ac) Um elemento de transposição de DNA classe 2, assim chamado por sua descobridora Barbara McClintock porque é necessário para ativar a quebra do cromossomo no locus de *Dissociação (Ds)*.

elemento classe 1 Um elemento de transposição que se move por um RNA intermediário. Também chamado de elemento RNA.

elemento classe 2 Um elemento de transposição que se move diretamente de um sítio no genoma para outro. Também chamado de elemento de DNA.

elemento curto intercalado (SINE) Um tipo de elemento de transposição classe 1 que não codifica a transcriptase reversa mas é tido como usando a transcriptase reversa codificada por LINE. *Veja também Alu.*

elemento de ação cis Um sítio na molécula de DNA (ou RNA) que funciona como um sítio de ligação para uma sequência específica de proteína de ligação ao DNA (ou RNA). O termo *ação cis* indica que a proteína ligada a esse sítio afeta apenas as sequências de DNA (ou RNA) vizinhas na mesma molécula.

elemento de Dissociação (Ds) Um elemento de transposição não-autônomo denominado por Barbara McClintock por sua habilidade em quebrar o cromossomo 9 de milho, mas apenas na presença de outro elemento chamado *Ativador (Ac)*.

elemento de sequência de inserção (IS) Uma peça móvel do DNA bacteriano (com tamanho de várias centenas de pares de nucleotídeos) capaz de inativar um gene no qual venha a inserir-se.

elemento de transposição Um termo geral para qualquer unidade genética que pode ser inserida em um cromossomo, ser excisada e ser reinserida em outra parte; inclui sequências de inserção e transposons.

elemento de transposição autônomo Um elemento de transposição que codifica a proteína, por exemplo, transposase ou transcriptase reversa, necessária para sua transposição e para a transposição de elementos não-autônomos na mesma família.

elemento DNA Um elemento de transposição classe 2 encontrado tanto em procariontes quanto em eucariontes, e assim denominado porque o elemento DNA participa diretamente da transposição.

elemento genético móvel *Veja elemento de transposição.*

elemento IS *Veja elemento de sequência de inserção.*

elemento não-autônomo Um elemento de transposição baseado nos produtos proteicos de elementos autônomos para sua mobilidade. *Dissociação (Ds)* é um exemplo de um elemento de transposição não-autônomo.

elemento nuclear longo intercalar (LINE) Um tipo de elemento de transposição classe I que codifica uma transcriptase reversa. LINE também são chamados retrotransposons não-LTR.

elemento P Um elemento de transposição de DNA em *Drosophila* que foi usado como uma ferramenta para mutagênese insercional e para transformação de linhagem germinativa.

elemento proximal do promotor A série de sítios de ligação do fator de transcrição situados perto do cerne do promotor.

elemento regulatório cis *Veja elemento de ação cis.*

elemento RNA Um elemento que se move através de um RNA intermediário. Também chamado de *elemento classe 1*.

elemento silenciador Uma sequência regulatória de ação cis que pode reduzir os níveis de transcrição de um promotor adjacente.

elemento tipo *copia* Um elemento de transposição (retrotransposon) de *Drosophila* que é flanqueado por longas repetições terminais e tipicamente codifica uma transcriptase reversa.

elemento *Ty* Um retrotransposon LTR de levedura; o primeiro isolado para qualquer organismo.

eletroforese Uma técnica para separação de componentes de uma mistura de moléculas (proteínas, DNA ou RNA) em um campo elétrico dentro de um gel.

eletroforese em gel Um método de separação molecular no qual o DNA, RNA ou proteínas são separados em uma matriz de gel de acordo com o tamanho molecular, com o uso de um campo elétrico para levar as moléculas através do gel em uma direção predeterminada.

embrióide Uma pequena massa em divisão de células monoplóides, produzidas de uma célula destinada a tornar-se uma célula de pólen pela exposição ao frio.

endogamia (1) Reprodução entre indivíduos dentro de um grupo ou subgrupo, em vez de aleatória em uma população. (2) Cruzamento entre parentes.

endogamia negativa Reprodução preferencial entre indivíduos que não são aparentados.

endogenoto *Veja merozigoto.*

endonuclease Uma enzima que corta as ligações fosfodiéster dentro de uma cadeia de nucleotídeos.

engenharia genética O processo de produzir DNA modificado em um tubo de ensaio e reintroduzir esse DNA nos organismos hospedeiros.

enzima Uma proteína que funciona como catalisadora.

enzima de restrição Uma endonuclease que irá reconhecer uma sequência-alvo de nucleotídeos específicos no DNA e quebrar a cadeia de DNA nesses pontos; é conhecida uma variedade dessas enzimas, sendo elas amplamente usadas na engenharia genética.

enzima processiva Como usada no Capítulo 7, descreve o comportamento da DNA polimerase III, que pode fazer centenas de rodadas de catálise sem se dissociar de seu substrato (o filamento de DNA-molde).

epigeneticamente silenciada A inativação reversível de um gene devida à estrutura da cromatina.

epistasia Uma situação na qual a expressão fenotípica diferencial de um genótipo em um locus depende do genótipo de outro locus; uma mutação que exerce sua expressão enquanto cancela a expressão dos alelos de outro gene.

epistasia de sinal A dependência da vantagem/desvantagem adaptativa de uma nova mutação em relação às mutações que já foram previamente fixadas.

equilíbrio de Hardy-Weinberg A distribuição estável de frequências dos genótipos A/A , A/a e a/a , na proporção de p^2 , $2pq$ e q^2 , respectivamente (onde p e q são as frequências dos alelos A e a), que é uma consequência de reprodução aleatória na ausência de mutação, migração, seleção natural ou deriva genética.

equilíbrio de ligação Distribuição de frequências de haplótipos em uma população tal que as frequências são os produtos aritméticos das frequências dos alelos em loci diferentes no haplótipo.

erros hereditários do metabolismo Distúrbios metabólicos hereditários, particularmente em relação a doença humana.

especiação O processo de formação de novas espécies pela divisão de uma espécie antiga em duas ou mais espécies novas incapazes de trocar genes uma com a outra.

especiação alopátrica A formação de novas espécies a partir de populações que são geograficamente isoladas umas das outras.

espécie Um grupo de organismos que podem trocar genes entre si mas são reprodutivamente isolados de outros grupos.

especificidade mutacional Uma constelação de danos mutacionais que caracterizam determinado mutágeno.

esporo (1) Em plantas e fungos, esporos sexuais são células haplóides produzidas por meiose. (2) Em fungos, os esporos assexuais são células somáticas que atuam como gametas ou unidades de dispersão.

esporófito A geração diplóide produtora de esporos sexuais no ciclo de vida das plantas, isto é, o estágio no qual ocorre a meiose.

EST *Veja sequência expressa marcada.*

estatística Uma quantidade característica computada de uma população, tal como a média.

estrutura primária de uma proteína A sequência de aminoácidos na cadeia polipeptídica.

estrutura quaternária de uma proteína A constituição multimérica de uma proteína.

estrutura secundária de uma proteína Uma disposição em espiral ou ziguezague da cadeia polipeptídica.

estrutura solenóide A disposição super-helicoidizada do DNA em cromossomos de núcleos eucarióticos que é produzida pela helicoidização de um colar contínuo de nucleossomos.

estrutura terciária de uma proteína O dobramento ou helicoidização da estrutura secundária para formar uma molécula globular.

estrutura teta (θ) Uma estrutura intermediária na replicação de um cromossomo bacteriano circular.

estruturas serialmente reiteradas Partes do corpo que são membros de séries repetidas, tais como dedos, costelas, dentes, membros e segmentos.

eucarionte Um organismo que tem células eucarióticas.

eucromatina Uma região cromossômica menos condensada, tida como contendo a maioria dos genes de funcionamento normal.

eugenia Reprodução humana controlada pelo estado em uma tentativa de melhorar as gerações futuras.

euplóide Uma célula tendo qualquer número de conjuntos cromossômicos completos de um organismo composto de tais células.

evento mutacional A ocorrência de uma mutação no tempo e no espaço.

evolução filética A formação de novas espécies relacionadas como resultado da divisão de espécies previamente existentes.

evolução neutra Mudanças evolucionárias não-adaptativas devidas a deriva genética aleatória.

excisar Descreve o que um elemento de transposição faz quando ele deixa um local cromossômico. Também chamado de transpor.

exconjugante Uma bactéria fêmea que acabou de estar em conjugação com um macho que contém um fragmento de DNA do macho.

exogamia forçada Evitação deliberada de reprodução entre parentes.

exogenoto *Veja merozigoto.*

éxon Qualquer trecho não-ínton da sequência codificante de um gene; juntos, os éxons correspondem ao mRNA que é traduzido em proteína.

exonuclease Uma enzima que corta nucleotídeos um de cada vez de uma ponta de uma cadeia polinucleotídica.

expansão de trinca A expansão de uma repetição de 3 pb de um número relativamente baixo de cópias para um alto número de cópias que é responsável por um número de doenças genéticas, tais como a síndrome do X frágil e doença de Huntington.

experimento de pulso-caça Um experimento no qual as células são cultivadas em um meio radioativo por um breve período (o pulso) e então transferidas para um meio não-radioativo por um período mais longo (a caça).

expressão ectópica A ocorrência da expressão gênica em um tecido no qual o gene normalmente não é expresso. Pode ser causada pela justaposição de novos elementos acentuadores no gene.

expressividade O grau no qual determinado genótipo é expresso no fenótipo.

fago *Veja bacteriófago.*

fago λ (lambda) Um tipo (espécie) de bacteriófago temperado.

fago recombinante A prole de fago de uma dupla infecção, levando alelos de ambas as linhagens infectantes.

fago temperado Um fago que pode tornar-se um profago.

fago virulento Um fago que não pode tornar-se um profago; uma infecção por tal fago sempre leva à lise da célula hospedeira.

família de genes Um conjunto de genes em um genoma, todos descendentes do mesmo gene ancestral.

fase G₀ A fase que pode ocorrer em G₁ da interfase do ciclo celular.

fase G₁ A parte da interfase do ciclo celular que precede a fase S.

fase G₂ A fase da interfase do ciclo celular que segue a fase S.

fase M A fase mitótica do ciclo celular.

fase S A parte da interfase do ciclo celular na qual ocorre a síntese de DNA.

fator 1 de montagem de cromatina (CAF-1) *Veja CAF-1.*

fator de ação trans Uma molécula regulatória difundível (quase sempre uma proteína) que se liga a um elemento de ação cis.

fator de alongamento G (EF-G) Uma proteína não-ribossômica que interage com o sítio A e promove a translocação do ribossomo.

fator de alongamento Tu (EF-Tu) Uma proteína não-ribossômica que se liga a um tRNA aminoacilado, formando um complexo ternário.

fator de fertilidade (fator F) Um episomo bacteriano cuja presença confere habilidade de doador (masculinidade).

fator de iniciação Uma proteína necessária para o início correto da tradução.

fator de liberação (RF) Uma proteína que se liga ao sítio A do ribossomo quando um códon de fim está no mRNA.

fator de transcrição Uma proteína que se liga a um elemento regulador de ação cis (p. ex., um acentuador) e, assim, direta ou indiretamente, afeta o início da transcrição.

fator F *Veja fator de fertilidade.*

fator F' Um fator de fertilidade ao qual foi incorporada uma parte do cromossomo bacteriano.

fator geral de transcrição (GTF) Um complexo proteico eucariótico que não toma parte na síntese de RNA, mas liga-se à região promotora para atrair e posicionar corretamente a RNA polimerase II para o início da transcrição.

fator R *Veja plasmídeo R.*

fator sigma (σ) Uma proteína bacteriana que, como parte da holoenzima RNA polimerase, reconhece as regiões -10 e -35 de promotores bacterianos, posicionando assim a holoenzima para iniciar corretamente a transcrição no sítio de início. O fator σ dissocia-se da holoenzima antes da síntese de RNA.

fenocópia Um fenótipo induzido experimentalmente que se assemelha ao fenótipo produzido por uma mutação.

fenótipo (1) A forma de alguma característica (ou grupo de características) em um indivíduo específico. (2) As manifestações externas detectáveis de um genótipo específico.

fenótipo dominante O fenótipo de um genótipo contendo o alelo dominante; o fenótipo parental que é expresso em um heterozigoto.

fenótipo instável Um fenótipo caracterizado pela freqüente reversão, seja somática e/ou germinativamente, devida à interação de elementos de transposição ao gene do hospedeiro.

fenótipo recessivo O fenótipo de um homozigoto para o alelo recessivo; o fenótipo parental que não é expresso em um heterozigoto.

feromônio Uma molécula que promove comportamento reprodutivo.

filamento codificante O filamento não-molde da molécula de DNA tendo a mesma seqüência que a do RNA transcrito.

filamento contínuo (leading) Na replicação do DNA, o filamento que é feito no sentido 5' para 3' pela polimerização contínua da ponta 3' em crescimento.

filamento de RNA anti-sentido Um filamento de RNA tendo uma seqüência complementar a um filamento de RNA transcrito.

filamento descontínuo (lagging) Na replicação do DNA, o filamento que é sintetizado aparentemente no sentido 3' para 5' pela ligação de fragmentos curtos sintetizados individualmente no sentido 5' para 3'.

filtragem de enriquecimento Uma técnica para recuperar mutantes auxotróficos em fungos filamentosos.

fingerprint O padrão característico de ponto produzido por eletroforese dos fragmentos polipeptídicos obtidos desnaturando determinada proteína com uma enzima proteolítica.

fingerprint de DNA O padrão de bandamento auto-radiográfico produzido quando o DNA é digerido com uma enzima de restrição que corta fora de uma família de VNTR (número variável de repetições em tandem) e uma transferência de Southern do gel de eletroforese é sondada com uma sonda específica de VNTR. Ao contrário dos verdadeiros fingerprints, esses padrões não são únicos para cada organismo individual.

FISH *Veja hibridização in situ com fluorescência.*

fissão binária Um processo no qual uma célula parental divide-se em duas células-filhas de tamanhos aproximadamente iguais.

fMet *Veja formilmetionina.*

forma ceto *Veja mudança tautomérica.*

forma enol *Veja mudança tautomérica.*

forma imino *Veja mudança tautomérica.*

formação de padrão Os processos de desenvolvimento que resultam nas complexas formas e estruturas dos organismos superiores.

formilmetionina (fMet) Um aminoácido especializado que é o primeiro a ser incorporado à cadeia polipeptídica na síntese de uma proteína.

forquilha de replicação O ponto no qual os dois filamentos de DNA são separados para permitir a replicação de cada filamento.

fosfato Um íon formado por quatro átomos de oxigênio ligados a um átomo de fósforo, ou o grupo químico formado pela ligação de um íon fosfato a outra espécie química por uma ligação éster.

fotoliase Uma enzima que corta dímeros de timina.

fragmento acêntrico Um fragmento cromossômico sem centrômero.

fragmento de Okazaki Um pequeno segmento de DNA unifilamentar sintetizado como parte do filamento de replicação descontínua na replicação do DNA.

fragmento de restrição Um fragmento de DNA que resulta de cortar o DNA com uma enzima de restrição.

freqüência alélica Uma medida do quanto um alelo é comum em uma população; a proporção de todos os alelos desse gene na população que são desse tipo específico.

freqüência de mutação A freqüência de mutantes em uma população.

freqüência de recombinantes (FR) A proporção (ou percentagem) de células ou indivíduos recombinantes.

freqüência gênica *Veja freqüência alélica.*

freqüência genotípica A proporção em uma população de membros com determinado genótipo.

freqüência relativa O número de ocorrências de alguma classe em uma população, expresso como uma proporção do total.

função de distribuição Um gráfico de alguma medida quantitativa precisa de uma característica em relação à sua freqüência de ocorrência.

função de mapa Uma fórmula que expressa a relação entre a distância em um mapa de ligação e a freqüência de recombinantes.

fusão gênica Um novo gene produzido pela justaposição de seqüências de DNA de dois genes separados. As fusões gênicas podem ser o resultado de rearranjos cromossômicos, inserção de elementos de transposição ou engenharia genética.

fuso Um conjunto de fibras microtubulares que parecem mover cromossomos eucarióticos durante a divisão.

fuso nuclear O conjunto de microtúbulos que se formam entre os pólos de uma célula durante a divisão nuclear; a função do fuso nuclear é separar os cromossomos ou cromátides para os pólos.

G *Veja guanina; guanosina.*

gama de hospedeiros O espectro de linhagens de determinada espécie bacteriana que pode ser infectada por determinada linhagem de fagos.

gameta Uma célula haplóide especializada que se funde com um gameta do sexo oposto, ou tipo reprodutivo, para formar um zigoto diplóide; nos mamíferos, um ovócito e um espermatozóide.

gametófito O estágio produtor de gametas haplóides no ciclo de vida das plantas; proeminente e independente em algumas espécies, mas reduzido ou parasita em outras.

gene A unidade fundamental física e funcional da hereditariedade, que leva informação de uma geração para a seguinte; um segmento de DNA

composto de uma região transcrita e uma sequência regulatória que possibilita a transcrição.

gene candidato Um gene seqüenciado de função previamente desconhecida que, devido à sua posição cromossômica ou alguma outra propriedade, torna-se candidato a determinada função, tal como determinação de doença.

gene de efeito materno Um gene que produz um efeito apenas quando presente na mãe.

gene de manutenção Um termo informal para um gene cujo produto é necessário em todas as células e desempenha uma função fisiológica básica.

gene de polaridade segmentar Em *Drosophila*, um membro de uma classe de genes que contribuem para os aspectos finais do estabelecimento do número correto de segmentos. As mutações de polaridade segmentar causam a perda ou mudança em uma parte comparável de cada um dos segmentos do corpo.

gene de regra dos pares Em *Drosophila*, um membro de uma classe de genes zigoticamente expressos que atuam em um estágio intermediário no processo de estabelecer os números corretos dos segmentos corpóreos. As mutações na regra dos pares têm metade do número normal de segmentos devido à perda de qualquer outro segmento.

gene endógeno Um gene que, normalmente, está presente em um organismo, em contraste com um gene exógeno de um organismo diferente que pode ser introduzido por técnicas transgênicas.

gene estrutural A parte de um gene codificante de sequência de aminoácidos de uma proteína.

gene gap Em *Drosophila*, uma classe de genes cardinais que são ativados no zigoto em resposta aos gradientes de posição ântero-posterior.

gene hemizigoto Um gene presente apenas em uma cópia em um organismo diplóide – por exemplo, um gene ligado ao X em um macho de mamífero.

gene homeótico Em animais superiores, um gene que controla o destino de segmentos ao longo do eixo ântero-posterior.

gene marcador Um gene em um locus bem conhecido cujos alelos diferentes podem ser distinguidos por algum fenótipo visível, útil para localizar regiões contendo QTL ligados ao gene marcador.

gene regulador Um gene cujo produto liga ou desliga a transcrição de genes estruturais.

gene repórter Um gene cuja expressão fenotípica é fácil de monitorar; usado para estudar as atividades histoespecíficas de promotor e acentuador nos transgenes.

gene silenciado Um gene que é reversivelmente inativado devido à estrutura da cromatina.

gene SRY O gene de masculinidade, que reside no cromossomo Y.

gene supressor tumoral Um gene que codifica uma proteína que suprime a formação de tumor. Os alelos tipo selvagem de genes supressores tumorais são tidos como funcionando como reguladores negativos da proliferação celular.

genes controlados coordenadamente Genes cujos produtos estão simultaneamente ativados ou reprimidos em paralelo.

genes Hox Membros dessa classe de genes são os que contêm grupos de homeoboxes, genes homeóticos que controlam a identidade das partes do corpo ao longo do eixo ântero-posterior da maioria dos animais bilaterais.

genes parálogos Dois genes na mesma espécie que evoluíram por duplicação gênica.

genética (1) O estudo dos genes. (2) O estudo da hereditariedade.

genética de populações O estudo das frequências de genótipos em populações e as mudanças nessas frequências que resultam dos padrões de reprodução, seleção natural, mutação, migração e chance aleatória.

genética de transmissão O estudo dos mecanismos para a passagem de um gene de uma geração para a seguinte.

genética direta O enfoque clássico da análise genética, no qual os genes são primeiro identificados por alelos mutantes e fenótipos mutantes, sendo depois clonados e submetidos a análise molecular.

genética molecular O estudo de processos moleculares subjacentes à estrutura e função gênica.

genética quantitativa O estudo da base genética da variação contínua no fenótipo.

genética química O uso de bibliotecas de pequenos ligandos químicos para interferir especificamente nas funções de proteínas individuais, fenocopiando assim os efeitos das mutações nos genes que codificam essas funções.

genética reversa Um procedimento experimental que começa com um segmento clonado de DNA ou sequência de proteína usada (através de mutagênese dirigida) para introduzir mutações programadas de volta ao genoma para investigar a função.

genoma O complemento inteiro de material genético em um conjunto cromossômico.

genoma aneuplóide Um genoma com número de cromossomos que difere do número normal dos cromossomos da espécie por um pequeno número de cromossomos.

Genômica A clonagem e caracterização molecular de genomas inteiros.

genômica comparativa Análise das relações das sequências genômicas de duas ou mais espécies.

genômica estrutural A análise em grande escala das estruturas proteicas tridimensionais.

genômica funcional O estudo dos padrões de expressão do transcrito e proteína e as interações moleculares a nível de genoma amplo.

genótipo A composição alélica específica de uma célula, seja de toda a célula ou, mais comumente, de certo gene ou conjunto de genes.

geração F₁ A primeira geração filial, produzida pelo cruzamento de duas linhagens parentais.

geração F₂ A segunda geração filial, produzida por autofecundação ou entrecruzamento da geração F₁.

geração parental As duas linhagens ou organismos individuais que constituem o começo de um experimento de cruzamento genético; sua prole constitui a geração F₁.

gerações filiais Gerações sucessivas de prole em uma série de cruzamentos controlados, começando com dois genitores específicos (a geração P) e autofecundando ou entrecruzando a prole de cada nova geração (F₁, F₂,...).

GGR Veja reparo genômico global.

GMO Veja organismo geneticamente modificado.

gradiente Uma mudança gradual em alguma propriedade quantitativa por uma distância específica.

grupo de ligação Um grupo de genes conhecido como estando ligado; um cromossomo.

GTF Veja fator geral de transcrição.

guanina (G) Uma base purínica que faz par com citosina.

guanosina (G) Um nucleosídeo contendo guanina como base.

haplóide Uma célula tendo um conjunto de cromossomos ou um organismo composto de tais células.

haplossuficiência Descreve um gene que, em uma célula diplóide, pode promover a função tipo selvagem apenas em uma cópia (dose).

haplótipo Uma classe genética descrita por uma sequência de DNA ou de genes que estão juntos no mesmo cromossomo físico.

helicase Uma enzima que rompe as pontes de hidrogênio no DNA e desenrola o DNA durante o movimento da forquilha de replicação.

helicoidização O alinhamento espontâneo de dois filamentos de DNA para formar uma dupla hélice.

helicoidização de filamento dependente de síntese (SDSA) Um mecanismo livre de erro para corrigir quebras bifilamentares que ocorrem após a replicação de uma região cromossômica em uma célula em divisão.

herança citoplasmática Herança por genes encontrados em organelas citoplasmáticas.

herança epigenética Modificações herdáveis no funcionamento gênico não devidas a mudanças na sequência de bases do DNA do organismo. São exemplos de herança epigenética a paramutação, inativação do cromossomo X e *imprinting* parental.

herança materna Um tipo de herança uniparental na qual toda a prole tem o genótipo e o fenótipo do genitor que age como fêmea.

herdabilidade Uma característica é herdável se a prole se assemelha aos genitores como resultado de similaridade genética.

herdabilidade ampla (H^2) A proporção da variância fenotípica total a nível populacional que é contribuída pela variância genética.

herdabilidade no sentido amplo (H^2) *Veja herdabilidade ampla.*

herdabilidade no sentido restrito (h^2) *Veja herdabilidade restrita.*

herdabilidade restrita (h^2) A proporção da variância fenotípica que pode ser atribuída à variância genética aditiva.

hereditariedade A similaridade biológica entre a prole e seus genitores.

heterocácion Uma cultura de células composta de dois tipos nucleares diferentes em um citoplasma comum.

heterocromatina Regiões cromossômicas condensadas densamente coradas, tidas como sendo na maior parte geneticamente inertes.

heterocromatina constitutiva Regiões cromossômicas de cromatina permanentemente condensada em geral ao redor de telômeros e centrômeros.

heterodímero Uma proteína que consiste em duas subunidades polipeptídicas não-idênticas.

heterodúplice Uma dupla hélice de DNA formada pela helicoidização de filamentos únicos de fontes diferentes; se existir uma diferença estrutural entre os filamentos, o heterodúplice pode apresentar anomalias tais como alças ou fivelas.

heteroplasmon Uma célula contendo uma mistura de citoplasmas geneticamente diferentes e, em geral, mitocôndrias diferentes ou cloroplastos diferentes.

heterozigose Uma medida da variação genética em uma população com relação a um locus, dada como a frequência de heterozigotos para esse locus.

heterozigoto Um organismo individual tendo um par de genes heterozigotos.

hexaplóide Uma célula tendo seis conjuntos cromossômicos ou um organismo composto de tais células.

Hfr *Veja célula com alta frequência de recombinação.*

hibridização *in situ* O uso de uma sonda marcada para se ligar a um cromossomo parcialmente desnaturado ou ao mRNA dentro dos tecidos.

hibridização *in situ* com fluorescência (FISH) Hibridização *in situ* com o uso de uma sonda acoplada a uma molécula fluorescente.

hibridizar (1) Formar um híbrido fazendo um cruzamento. (2) Helicoidizar filamentos complementares de ácidos nucleicos de fontes diferentes.

híbrido (1) Um heterozigoto. (2) Uma prole individual de qualquer cruzamento entre genitores de genótipos diferentes.

hifa Uma estrutura filamentar (composta de células ligadas ponta a ponta) que forma o principal tecido em muitas espécies de fungos.

hiperacetilação Uma abundância de grupos acetil ligados a alguns aminoácidos das caudas de histonas. A cromatina transicionalmente ativa é geralmente hiperacetilada.

hipoacetilação Poucos grupos acetil em alguns aminoácidos das caudas de histona. A cromatina transicionalmente inativa é geralmente hipoacetilada.

hipótese de fator múltiplo Uma hipótese que explica a variação quantitativa supondo a interação de um grande número de genes (poligenes), cada um com um pequeno efeito aditivo na característica.

hipótese nula Uma hipótese que propõe ausência de diferença entre dois ou mais grupos de dados.

hipótese um gene-um polipeptídeo Uma hipótese da metade do século XX originalmente propondo que cada gene (seqüência de nucleotídeos) codificava uma seqüência de polipeptídeos; geralmente verdadeira, com a exceção de RNA funcional não-traduzido.

hipótese um gene-uma enzima A hipótese proposta por George Beadle e Edward Tatum, subsequentemente refinada para *hipótese um gene-um polipeptídeo*.

histograma de frequência Uma “curva íngreme” na qual as frequências de várias classes arbitrariamente delimitadas são grafadas.

histona Um tipo de proteína básica que forma a unidade ao redor da qual o DNA é helicoidizado nos nucleossomos dos cromossomos eucarióticos.

histona desacetilase (HDAT) A atividade enzimática que remove um grupo acetil de uma cauda de histona, que promove a repressão da transcrição gênica.

holoenzima DNA polimerase III (holoenzima DNA pol III) Em *E. coli*, o grande complexo multissubunitário na forquilha de replicação que consiste em dois cernes catalíticos e muitas proteínas acessórias.

holoenzima pol III *Veja holoenzima DNA polimerase III.*

homeobox (boxe homeótico) Uma família de seqüências de DNA com 180 pb bem similares que codificam uma seqüência polipeptídica chamada de homeodomínio; uma seqüência de ligação ao DNA específica de seqüência. Embora o homeobox tenha sido primeiro descoberto em todos os genes homeóticos, hoje sabemos que ele codifica um motivo de ligação ao DNA muito mais disperso.

homeodomínio Uma família de seqüências altamente conservada, com 60 aminoácidos de tamanho e encontrada dentro de um grande número de fatores de transcrição, que podem formar estruturas hélice-giro-hélice e ligar-se ao DNA de modo específico para seqüência.

homeose A substituição de uma parte do corpo por outra. A homeose pode ser causada por fatores ambientais que levam a anomalias de desenvolvimento ou por mutações.

homólogo Um membro de um par de cromossomos homólogos.

homozigose Refere-se ao estado de levar um par de alelos idênticos em um locus.

homozigose por descendência Homozigose que resulta da herança de duas cópias de um gene que estava presente em um ancestral.

homozigoto Um organismo individual que é homozigoto.

homozigoto dominante Refere-se a um genótipo tal como A/A.

homozigoto recessivo Refere-se a um genótipo tal como a/a.

hormônio Uma molécula que é secretada por um órgão endócrino no sistema circulatório e atua como uma molécula sinalizadora de longo alcance ativando receptores nas células-alvo.

identidade segmentar O processo pelo qual são estabelecidos os segmentos dos destinos anterior-posterior.

imprinting genômico Um fenômeno no qual um gene herdado de um dos genitores não é expresso, muito embora ambas as cópias gênicas sejam funcionais. Os genes imprintados são metilados e inativados na formação de gametas masculinos ou femininos.

imprinting materno A expressão de um gene apenas quando herdado do pai, pois a cópia do gene herdada da mãe é inativa devido à metilação no curso da formação de gametas.

imprinting parental Um fenômeno epigenético no qual a atividade de um alelo depende de ter sido ele herdado do pai ou da mãe. Se for herdado do pai é inativo, devido à metilação no curso da formação de gametas; se herdado da mãe, ocorre a expressão.

in vitro Em uma situação experimental, fora do organismo (literalmente, “no vidro”).

in vivo Em uma célula viva ou organismo.

inativação do cromossomo X O processo pelo qual os genes em um cromossomo X em um mamífero podem ser completamente reprimidos como parte do mecanismo de compensação de dose. *Veja também compensação de dose; corpúsculo de Barr.*

indução (1) O alívio da repressão de um gene ou conjunto de genes sob controle negativo. (2) Uma interação de duas ou mais células ou tecidos que é necessária para que uma dessas células ou tecidos mude seu destino desenvolvimental.

indução zigótica A súbita liberação de um fago lisogênico de um cromossomo Hfr quando o profago entra na célula F⁻ seguida da lise subsequente da célula receptora.

indutor Um agente ambiental que dispara a transcrição de um óperon.

infecção dupla Infecção de uma bactéria com dois fagos geneticamente diferentes.

infecção mista A infecção de uma cultura bacteriana com dois genótipos de fagos diferentes.

informação posicional O processo pelo qual sinais químicos que estabelecem o destino celular ao longo do eixo geográfico são estabelecidos em um embrião em desenvolvimento ou primórdio de tecido.

iniciação O primeiro estágio da transcrição ou tradução. Sua principal função, na transcrição, é posicionar corretamente a RNA polimerase antes do estágio de alongamento e, na tradução, é posicionar corretamente o primeiro aminoacil-tRNA no sítio P.

iniciador Um tRNA especial que insere o primeiro aminoácido de uma cadeia polipeptídica no sítio P do ribossomo no começo da tradução. O aminoácido levado pelo iniciador na bactéria é *N*-formilmetionina.

insuladores de bloqueio de acentuador Elementos regulatórios posicionados entre um promotor e um acentuador. Sua presença impede que o promotor seja ativado pelo acentuador.

integração ectópica Em um organismo transgênico, a inserção de um gene introduzido em um sítio que não seu locus usual.

interação gênica A colaboração de vários genes diferentes na produção de um fenótipo característico (ou grupo correlato de características).

interação ligando-receptor A interação entre uma molécula (geralmente de origem extracelular) e uma proteína em uma célula-alvo. Um tipo de interação ligando-receptor pode ser entre hormônios esteróides e seus receptores citoplasmáticos ou nucleares. Outra pode ser entre ligandos polipeptídicos secretados e receptores transmembranares.

interações combinatórias Interações entre combinações diferentes de fatores de transcrição que resultam em padrões distintos de expressão gênica.

interatoma Todo o conjunto de interações moleculares dentro das células, incluindo em particular interações proteína-proteína, proteína-RNA e proteína-DNA.

intérfase O estágio do ciclo celular entre as divisões nucleares, quando os cromossomos estão distendidos e funcionalmente ativos.

interferência Uma medida da independência dos crossings uns dos outros, calculada subtraindo o coeficiente de coincidência de 1.

interferência de cromátide Uma situação na qual a ocorrência de um crossing entre quaisquer duas cromátides não-irmãs pode ser mostrada afetando a probabilidade de essas cromátides participarem de outros crossings na mesma meiose.

íntron *Veja sequência intercalar.*

íntron auto-eliminável O primeiro exemplo de RNA catalítico; nesse caso, um íntron que pode ser removido de um transcrito sem a ajuda de uma enzima.

inversão Uma mutação cromossômica que consiste na remoção de um segmento cromossômico, sua rotação em 180° e sua reinserção no mesmo local.

inversão heterozigota Um diplóide com um homólogo normal e um invertido.

inversão paracêntrica Uma inversão não incluindo o centrômero.

inversão pericêntrica Uma inversão que inclui o centrômero.

isoformas Relacionadas por proteínas diferentes. Elas podem ser geradas por recomposição alternativa de um gene.

isolamento pós-zigótico A falha em duas espécies de trocarem genes porque os zigotos formados pela reprodução entre eles são inviáveis ou inférteis.

isolamento pré-zigótico A falha em duas espécies de trocarem genes devido a uma barreira ou para seu cruzamento ou para a fertilização de gametas de uma espécie por gametas de outra espécie.

isótopo Uma das várias formas de um átomo com mesmo número atômico mas massas atômicas diferentes.

junção de Holliday Uma estrutura de DNA lembrando um crossing unifilamentar, produzida durante a recombinação meiótica entre as cromátides não-irmãs de um par de homólogos.

junção de pontas não-homólogas (NHEJ) Um mecanismo usado por eucariontes para reparar quebras bifilamentares.

junção dupla de Holliday Os dois crossings unifilamentares (um só é chamado uma junção de Holliday) produzidos durante a recombinação homóloga, de acordo com o modelo de quebra bifilamentar. O modelo propõe que essas estruturas instáveis sejam resolvidas para produzir um crossing recíproco bifilamentar.

kb *Veja quilobase.*

λgal Um fago λ que leva um gene *gal* bacteriano e defeituoso (*d*) para alguma função do fago.

lei da segregação igual A produção de números iguais (50%) de cada alelo nos produtos meióticos (p. ex., gametas) de um meiócito heterozigoto.

leituras de pontas pareadas Na montagem de uma sequência *shotgun* de genoma inteiro, as sequências de DNA que correspondem a ambas as pontas de um inserto de DNA genômico em um clone recombinante.

lesão Uma área danificada em um gene (sítio mutante), cromossomo ou proteína.

lesão espontânea Dano ao DNA que ocorre na ausência de exposição a mutágenos; devida primariamente à ação mutagênica de subprodutos do metabolismo celular.

letal sintético Refere-se a um mutante duplo que é letal, enquanto as mutações componentes isoladas não o são.

ligação A associação de genes no mesmo cromossomo.

ligação ao sexo A localização de um gene em um cromossomo sexual.

ligação ao X O padrão de herança de genes encontrados no cromossomo X mas não no cromossomo Y.

ligação ao X e Y O padrão de herança de genes encontrados tanto no cromossomo X quanto no Y (raro).

ligação ao Y O padrão de herança de genes encontrados no cromossomo Y mas não no cromossomo X (raro).

ligação fosfodiéster Uma ligação entre um grupo açúcar e um grupo fosfato; tais ligações formam o arcabouço açúcar-fosfato do DNA.

ligação peptídica Uma ligação entre dois aminoácidos.

ligados Situação na qual dois genes estão no mesmo cromossomo conforme deduzido pelas frequências de recombinantes menores que 50%.

ligando *Veja interação ligando-receptor.*

ligase Uma enzima que pode unir uma ligação fosfodiéster quebrada em um ácido nucleico.

LINE *Veja elemento nuclear longo intercalar.*

linha de regressão de y em x A linha reta que passa através de um grupo de pontos que relaciona uma variável *y* com uma variável *x* e minimiza os desvios dos pontos da linha na direção *y*.

linha de regressão dos menores quadrados A linha reta que passa por um grupo de pontos relacionando uma variável *x* com uma variável *y* de modo que a soma dos desvios quadrados de todos os valores *y* de uma linha seja tão pequena quanto possível.

linhagem Uma linhagem de reprodução pura, geralmente de organismos haplóides, bactérias ou vírus.

linhagem de cruzamento puro Um grupo de organismos individuais idênticos que sempre produzem prole do mesmo fenótipo quando entrecruzados.

linhagem germinativa A linhagem celular em um eucarionte com vários tecidos da qual são derivados os gametas.

linhagem pura (1) Uma população de indivíduos todos portando o genótipo totalmente homozigoto idêntico. (2) *Veja linhagem de cruzamento puro.*

linhagem segregante Um grupo de indivíduos geneticamente variantes que são a prole de uma população híbrida resultante do crossing de duas linhagens diferentes geneticamente uniformes.

lise A ruptura e morte de uma bactéria na liberação de uma prole de fagos.

lisogenia *Veja bactéria lisogênica.*

locus (plural **loci**) *Veja locus gênico.*

locus de característica quantitativa (QTL) Um gene que afeta a variação fenotípica em características de variação contínua tais como altura e peso, entre outras.

locus gênico O local específico em um cromossomo onde um gene está situado.

longa repetição terminal (LTR) Uma repetição direta de sequência de DNA nas pontas 5' e 3' de retrovírus e retrotransposons.

LTR *Veja longa repetição terminal.*

macromolécula Um grande polímero tal como DNA, uma proteína ou um polissacarídeo.

mapa cromossômico Uma representação de todos os cromossomos no genoma como linhas, marcadas com as posições dos genes conhecidos por seus fenótipos mutantes mais os marcadores moleculares. Baseado na análise da frequência de recombinação.

mapa de ligação Um mapa cromossômico; um mapa abstrato de loci cromossômicos que é baseado nas frequências de recombinantes.

mapa de recombinação Um mapa cromossômico no qual as posições dos loci mostrados são baseadas nas frequências de recombinantes.

mapa de restrição Um mapa de uma região cromossômica mostrando as posições de sítios-alvo de uma ou mais enzimas de restrição.

mapa de sintenia Um diagrama mostrando a localização de trechos contínuos do genoma de uma espécie no genoma de outra espécie.

mapa físico O mapa ordenado e orientado de fragmentos de DNA clonados de um genoma.

mapa gênico Uma estrutura linear de sítios mutantes dentro de um gene, com base nas várias frequências de recombinação interalélica (intra-gênica).

mapeamento de deleção O uso de um conjunto de deleções conhecidas para mapear novas mutações recessivas por pseudodominância.

mapeamento de RFLP Uma técnica pela qual os polimorfismos de comprimento de fragmentos de restrição são usados como loci de referência para mapeamento em relação a genes conhecidos ou outros loci de RFLP.

mapeamento físico Mapeamento das posições de fragmentos genômicos clonados.

máquina de dobramento chaperonina Um complexo proteico multisubunitário no qual uma proteína recém-sintetizada dobra-se em sua conformação ativa.

marca epigenética Uma alteração herdável, tal como a metilação do DNA ou uma modificação de histona, que deixa inalterada a sequência de DNA.

marcação *Vea marcação gênica.*

marcação de transposon Um método usado para identificar e isolar um gene hospedeiro por meio da inserção de um elemento de transposição clonado no gene.

marcação gênica O uso de um trecho de DNA exógeno ou um transposon para marcar um gene, de modo que um clone desse gene possa ser prontamente identificado em uma biblioteca.

marcador *Vea marcador genético.*

marcador genético Um alelo usado como sonda experimental para marcar um organismo individual, um tecido, uma célula, um núcleo, um cromossomo ou um gene.

marcador microssatélite Uma diferença no DNA no mesmo locus em dois genomas que é devida a tamanhos diferentes de repetições de um microssatélite.

marcador minissatélite Locus heterozigoto representando um número variável de repetições em tandem de uma unidade com 15 a 100 nucleotídeos de tamanho.

marcador molecular Um sítio de heterozigose de DNA, não necessariamente associado a variação fenotípica, usado como marcação para determinado locus cromossômico.

marcador não-selecionado Em um experimento de recombinação bacteriana, um alelo na prole para a frequência de sua co-segregação com um alelo selecionado ligado.

matriz de leitura Sequência de códons determinada pelos nucleotídeos de leitura em grupos de três a partir de um códon de início.

matriz de leitura aberta (ORF) Um trecho do tamanho de um gene de uma sequência de DNA que inicia com um códon de começo e termina com um códon de fim; é suposta como sendo a sequência codificante de um gene.

Mb *Vea megabase.*

mecanismo dependente de rô Um dos dois mecanismos usados para terminar a transcrição bacteriana.

mecanismo intrínseco Um dos dois mecanismos principais para o término da transcrição em bactérias.

média A média aritmética.

megabase (Mb) Um milhão de pares de nucleotídeos.

meio Qualquer material no qual são cultivadas culturas experimentais.

meio mínimo Meio contendo apenas sais inorgânicos, uma fonte de carbono e água.

meiócito Uma célula na qual ocorre meiose.

meiose Duas divisões nucleares sucessivas (com as correspondentes divisões celulares) que produzem gametas (em animais) ou esporos sexuais (em plantas e fungos) tendo metade do material genético da célula original.

meiosporo Uma célula que é um dos produtos da meiose nas plantas.

melanina Um pigmento marrom ou preto encontrado em grande parte do reino animal.

melanismo A ocorrência de formas escuras de uma população ou espécie.

merozigoto Uma *E. coli* parcialmente diplóide formada de um cromossomo completo (o endogenoto) mais um fragmento (o exogenoto).

metabolismo As reações químicas que ocorrem em uma célula viva.

metáfase Um estágio intermediário da divisão nuclear no qual os cromossomos alinham-se no plano equatorial da célula.

metilação A modificação de uma molécula pela adição de grupo metil.

método híbrido endogâmico Um método de cruzamento de plantas que produz grandes números de plantas heterozigotas geneticamente idênticas cruzando linhagens diferentes que foram tornadas homozigotas por muitas gerações de endogamia.

microarranjo Um conjunto de DNA contendo todos os genes ou a maioria deles em um genoma depositado em um pequeno chip de vidro.

microRNA (miRNA) Uma classe de RNA funcional que regula a quantidade de proteína produzida por um gene eucariótico.

microtúbulo Parte do sistema de cabos de diâmetro maior do citoesqueleto. Um microtúbulo é composto de subunidades de tubulina polimerizadas formando um tubo oco.

miRNA *Vea microRNA.*

mitose Um tipo de divisão nuclear (de ocorrência na divisão celular) que produz dois núcleos-filhos idênticos ao núcleo parental.

moda A classe única em uma distribuição estatística que tem a maior frequência.

modelo de decisão precoce de crossing Um modelo proposto para explicar o mecanismo de recombinação homóloga. O modelo diz que duas vias distintas são responsáveis pelos resultados crossing ou não-crossing da recombinação meiótica.

modelo de DNA heterodúplice Um modelo que explica tanto o crossing over quanto a conversão gênica, supondo a produção de um trecho curto de DNA heterodúplice (formado por ambos os DNA parentais) na vizinhança de um quiasma.

molde Um “molde” molecular que dá forma à estrutura ou sequência de outra molécula; por exemplo, a sequência de nucleotídeos do DNA atua como um molde para controlar a sequência de nucleotídeos do RNA durante a transcrição.

molécula-filha Um dos dois produtos da replicação do DNA; composta de um filamento-molde e um filamento recém-sintetizado.

monóibrido Um heterozigoto em um só locus do tipo *A/a*.

monoplóide Uma célula tendo apenas um conjunto cromossômico (geralmente como uma aberração) ou um organismo composto de tais células.

monossômico Uma célula ou organismo individual que é basicamente diplóide mas tem apenas uma cópia de determinado tipo cromossômico e, portanto, o número cromossômico $2n + 1$. (7 n - 3)

montagem de sequência A compilação de milhares ou milhões de sequências independentes de DNA em um conjunto de contigs e arcaúços.

morfo Uma forma de um polimorfismo genético; o morfo pode ser ou um fenótipo ou uma sequência molecular.

morfógeno Uma molécula que pode induzir a aquisição de diferentes destinos celulares ao nível do morfógeno ao qual a célula é exposta.

morte celular programada *Vea apoptose.*

motivo Uma sequência curta de DNA associada a determinado papel funcional.

mRNA (RNA mensageiro) Uma molécula de RNA transcrita do DNA de um gene; uma proteína é traduzida dessa molécula de RNA pela ação dos ribossomos.

mRNA policistrônico Um mRNA de óperon que codifica mais de uma proteína.

mtDNA DNA mitocondrial.

mudança de matriz de leitura A inserção ou deleção de um par ou pares de nucleotídeos, causando uma perturbação da matriz de leitura traducional.

mudança genética Um segmento de DNA regulador e a(s) proteína(s) reguladora(s) que se liga(m) a ele e governa(m) o estado transcricional de um gene ou conjunto de genes.

mudança tautomérica A isomerização espontânea de uma base nitrogenada de sua forma normal ceto para uma alternativa enol (ou imino) de ligação ao hidrogênio.

multímero Uma proteína que consiste em duas ou mais subunidades.

mundo de RNA O nome de uma teoria popular segundo a qual o RNA deve ter sido o material genético das primeiras células porque apenas o RNA é conhecido como codificando a informação genética e catalisando reações biológicas.

mutação (1) O processo que produz um gene ou um conjunto de cromossomos diferente do tipo selvagem. (2) O gene ou conjunto cromossômico que resulta de tal processo.

mutação condicional Uma mutação que tem o fenótipo tipo selvagem sob certas condições ambientais (permissivas) e um fenótipo mutante sob outras condições (restritivas).

mutação cromossômica Qualquer tipo de mudança na estrutura ou número de cromossomos.

mutação de acentuador Uma mutação modificadora que torna o fenótipo de uma mutação de outro gene mais grave.

mutação de ganho de função Uma mutação que resulta em uma nova habilidade funcional para uma proteína, detectável a nível fenotípico. As mutações de ganho de função em geral são dominantes, pois apenas uma cópia do novo gene é necessária para produzir a nova função.

mutação de perda de função Uma mutação que elimina parcial ou totalmente a atividade gênica normal. As mutações hipomórficas e nulas são mutações de perda de função.

mutação de ponto Uma pequena lesão, geralmente a inserção ou deleção de um único par de bases.

mutação de sentido trocado Substituição de par de nucleotídeos dentro de uma região codificante de proteína que leva à substituição de um aminoácido por outro aminoácido.

mutação direta Uma mutação que converte um alelo tipo selvagem em um alelo mutante.

mutação espontânea Uma mutação que ocorre na ausência de exposição a mutágenos.

mutação gênica Uma mudança no DNA que está totalmente contida dentro de um único gene, tal como uma *mutação de ponto* ou uma *mutação indel* de vários pares de nucleotídeos.

mutação germinativa Mutações nas células que são destinadas a se desenvolver em gametas.

mutação homeótica Uma mutação que leva à substituição de uma parte do corpo por outra.

mutação indel Mutação na qual um ou mais pares de nucleotídeos são adicionados ou deletados.

mutação induzida Uma mutação que surge pela ação de um agente que aumenta a taxa na qual ocorre uma mutação.

mutação instável Uma mutação que tem alta frequência de reversão; uma mutação causada pela inserção de um elemento controlador cuja saída subsequente produz uma reversão.

mutação letal *Veja alelo letal.*

mutação modificadora Uma mutação em um locus que afeta a expressão fenotípica de um alelo mutante em outro locus.

mutação morfológica Uma mutação que afeta alguns dos aspectos de um organismo individual.

mutação neutra Uma mutação que não tem efeito na sobrevivência e taxa reprodutiva dos organismos que a possuem.

mutação nula Uma mutação que resulta na ausência completa de função do gene.

mutação pleiotrópica Uma mutação que afeta várias características diferentes.

mutação polar Uma mutação que afeta a transcrição ou tradução de parte do gene ou óperon apenas em um lado do sítio mutante. Os exemplos são mutações sem sentido, mutações de mudança de matriz de leitura e mutações induzidas IS.

mutação por deleção *Veja mutação indel.*

mutação reversa *Veja reversão.*

mutação sem sentido Substituição de um par de nucleotídeos dentro de uma região codificante de proteína que muda um códon para um aminoácido por um códon de término (sem sentido).

mutação sensível à temperatura (ts) Uma mutação condicional que produz o fenótipo mutante em uma faixa de temperatura e o fenótipo tipo selvagem em outra faixa de temperatura.

mutação silenciosa Uma mutação que não tem efeito na função de um produto gênico.

mutação sinônima Uma mutação que muda um códon para um aminoácido por outro códon para o mesmo aminoácido. Também chamada de *mutação silenciosa*.

mutação somática Uma mutação que surge em uma célula somática e, conseqüentemente, não é passada pela linhagem germinativa para a geração seguinte.

mutação vazante (leaky) Uma mutação que confere um fenótipo mutante mas ainda conserva um nível baixo, porém detectável, da função tipo selvagem.

mutagênese Um experimento no qual organismos experimentais são tratados com um mutágeno, e sua prole é examinada quanto a fenótipos mutantes específicos.

mutagênese dirigida Alteração de alguma parte específica de um gene clonado e reintrodução do gene modificado de volta ao organismo.

mutagênese in vitro A produção de mutações aleatórias ou específicas em um pedaço de DNA clonado. Tipicamente, o DNA será reembalado e introduzido em uma célula ou organismo para avaliar os resultados da mutagênese.

mutagênese insercional Uma mutação que surge pela interrupção de um gene por um DNA exógeno, tal como uma construção transgênica ou um elemento de transposição.

mutagênese sítio-dirigida A alteração de uma parte específica de um segmento clonado de DNA seguido da reintrodução do DNA modificado de volta ao organismo para avaliar o fenótipo mutante ou a produção da proteína mutante.

mutágeno Um agente capaz de aumentar a taxa de mutação.

mutante Um organismo ou célula portando uma mutação.

mutante duplo Genótipo com alelos mutantes de dois genes diferentes.

mutante resistente Um mutante que pode crescer em um ambiente normalmente tóxico.

mutante vazante (leaky) Um mutante (tipicamente, um auxotrófico) que resulta de uma inativação parcial, ao invés de completa, da função tipo selvagem.

n O número haplóide; o número de cromossomos em um genoma.

NAHR *Veja recombinação homóloga não-alélica.*

não-ativa Refere-se a uma proteína que é dobrada incorretamente.

não-disjunção A falha de homólogos (na meiose) ou cromátides-irmãs (na mitose) em se separar apropriadamente para os pólos opostos.

nascente Recém-sintetizada; usada para descrever proteínas.

nativa Refere-se a uma proteína que é dobrada corretamente.

Neurospora Um fungo rosa, comumente encontrado crescendo em alimentos velhos.

NHEJ *Veja junção de pontas não-homólogas.*

NLS *Veja sequência de localização nuclear.*

nocaute Inativação de um gene específico. O mesmo que *perturbação gênica*.

nocaute gênico A inativação de um gene por uma mutação de ocorrência natural ou pela integração de um fragmento de DNA especialmente construído e introduzido. Em alguns sistemas, tal inativação é aleatória, com o uso de construções transgênicas que se inserem em muitos locais diferentes no genoma. Em outros sistemas, isso pode ser feito de modo direto. *Veja também nocaute gênico direcionado.*

nocaute gênico direcionado A introdução de uma mutação nula em um gene por uma alteração programada em uma sequência de DNA clonado que é então introduzida no genoma por recombinação homóloga e substituição do alelo normal.

norma de reação O padrão de fenótipos sob condições ambientais diferentes.

NPD *Veja ditipo não-parental.*

nuclease Uma enzima que pode degradar o DNA quebrando suas ligações fosfodiéster.

nucleóide Uma massa de DNA dentro de um cloroplasto ou mitocôndrias.

nucléolo Uma organela encontrada no núcleo, contendo rRNA e várias cópias dos genes codificando rRNA.

nucleosídeo Uma base nitrogenada ligada a uma molécula de açúcar.

nucleossomo A unidade básica da estrutura cromossômica eucariótica; uma bola com oito moléculas de histona que é envolvida por duas voltas de DNA.

nucleotídeo Uma molécula composta de uma base nitrogenada, um açúcar e um grupo fosfato; o bloco estrutural dos ácidos nucleicos.

nulissômico Refere-se a uma célula ou organismo individual com a falta de um tipo de cromossomo, com um número cromossômico tal como $n - 1$ ou $2n - 2$.

número haplóide O número de cromossomos no conjunto genômico básico de uma espécie.

número variável de repetições em tandem (VNTR) Um locus cromossômico no qual uma sequência particular repetida está presente em números diferentes em indivíduos diferentes ou em dois homólogos diferentes em um indivíduo diplóide.

O *Veja origem de replicação.*

óctade Um asco contendo oito ascósporos e produzido em espécies nas quais a tétrade normalmente sofre uma divisão mitótica pós-meiótica.

oligonucleotídeo Um curto segmento de DNA sintético.

oncogene Uma mutação de ganho de função que contribui para a produção de um câncer.

oncoproteína O produto proteico de uma mutação oncogênica.

operador Uma região do DNA em uma extremidade de um óperon que atua como sítio de ligação para uma proteína repressora.

óperon Um conjunto de genes estruturais adjacentes cujo mRNA é sintetizado em um trecho, mais os sinais regulatórios adjacentes que afetam a transcrição de genes estruturais.

ORC *Veja complexo de reconhecimento da origem.*

ORF *Veja matriz de leitura aberta.*

organela Uma estrutura subcelular tendo uma função especializada – por exemplo, a mitocôndria, o cloroplasto ou o fuso acromático.

organismo doador Um organismo que fornece DNA para uso na tecnologia de DNA recombinante ou transformação mediada por DNA.

organismo geneticamente modificado (GMO) Termo popular para um organismo transgênico, especialmente aplicado a organismos agrícolas transgênicos.

organismo-modelo Uma espécie escolhida para uso em estudos de genética porque é bem adequada ao estudo de um ou mais processos genéticos.

organismo transgênico Um organismo cujo genoma foi modificado pela aplicação externa de novo DNA.

organizador nucleolar Uma região (ou regiões) do conjunto cromossômico que está fisicamente associada ao nucléolo e contém genes de rRNA.

oriC A origem fixada de replicação em um cromossomo de *E. coli*.

orientação em polaridade inversa Termo usado para descrever a orientação oposta dos dois filamentos da dupla hélice de DNA; a ponta 5' de um filamento fica alinhada com a ponta 3' do outro filamento.

origem (O) *Veja origem de replicação.*

origem de replicação (O) O ponto de uma sequência específica no qual é iniciada a replicação do DNA.

ortólogos Genes em espécies diferentes que evoluíram de um gene ancestral comum por especiação.

oscilação (wobble) A habilidade de algumas bases na terceira posição de um anticódon no tRNA para formar pontes de hidrogênio de vários modos, causando o alinhamento com vários códons diferentes possíveis.

PAC (cromossomo artificial baseado em P1) Um derivado do fago P1 construído como um vetor de clonagem para levar grandes insertos.

padrão de segregação de primeira divisão (padrão M_I) Um padrão linear de fenótipos de esporos dentro de um asco para um par de alelos particular, produzido quando os alelos vão para núcleos separados na primeira divisão de meiose, mostrando que não ocorreu nenhum crossing entre o par de alelos e o centrômero.

padrão de segregação de segunda divisão (padrão M_{II}) Um padrão de genótipos de ascósporos para um par de genes mostrando que dois alelos se separam em dois núcleos diferentes apenas na segunda divisão meiótica, como resultado de um crossing entre esse par de genes e seu centrômero; só pode ser detectado em um asco linear.

padrão M_I *Veja padrão de segregação de primeira divisão.*

padrão M_{II} *Veja padrão de segregação de segunda divisão.*

palíndromo Uma sequência de DNA que tem uma simetria rotacional de 180°.

panorama adaptativo Uma superfície plotada em um gráfico tridimensional com todas as combinações possíveis de frequências alélicas para loci diferentes plotados no plano e com a adaptabilidade média para cada combinação plotada em terceira dimensão.

par de genes As duas cópias de um tipo particular de gene presentes em uma célula diplóide (uma em cada conjunto de cromossomos).

par de genes heterozigotos Um par de genes tendo alelos diferentes nos dois conjuntos cromossômicos do indivíduo diplóide, por exemplo, A/a ou A^1/A^2 .

par de nucleotídeo Um par de nucleotídeos (um em cada filamento de DNA) que são ligados por pontes de hidrogênio.

paradoxo do valor C A discrepância (ou falta de correlação) entre o conteúdo de DNA de um organismo e sua complexidade biológica.

parálogos Genes que estão relacionados por duplicação gênica em um genoma.

partenogênese A produção de prole por uma fêmea sem contribuição genética do macho.

PCNA *Veja antígeno nuclear de proliferação celular.*

PCR *Veja reação em cadeia da polimerase.*

PD *Veja ditipo parental.*

pedigree Uma “árvore genealógica”, desenhada com símbolos genéticos padrão, mostrando os padrões de herança para características fenotípicas específicas.

penetrância A proporção de indivíduos com um genótipo específico que manifesta esse genótipo em âmbito fenotípico.

pentaplóide Um organismo individual com cinco conjuntos de cromossomos.

peptídeo *Veja aminoácido.*

pequeno RNA de interferência (siRNA) Pequenos RNA bifilamentares produzidos pela clivagem de longos RNA bifilamentares por Dicer.

pequeno RNA nuclear (snRNA) Qualquer um dos vários pequenos RNA encontrados no núcleo eucariótico, onde ajudam nos eventos de processamento de RNA.

perturbação gênica A inativação de um gene pela integração de um fragmento de DNA introduzido especialmente construído.

PEV *Veja variação de efeito de posição.*

PIC *Veja complexo de pré-iniciação.*

pico adaptativo Um ponto alto (talvez um de vários) em um panorama adaptativo. A seleção tende a levar a composição genotípica da população para um genótipo correspondente a um pico adaptativo.

piebaldismo Um fenótipo de mamífero no qual manchas na pele não são pigmentadas devido à falta de melanócitos; geralmente herdado como autossômico dominante.

pilus (plural **pili**) Um tubo de conjugação; um apêndice oco tipo pêlo de uma célula doadora *E. coli* que atua como uma ponte para a transmissão de DNA doador para a célula receptora durante a conjugação.

“pintar” cromossomos Uso de uma bateria de sondas de hibridização *in situ* marcadas fluorescentemente de modo que tais sondas, para cada cromossomo, contenham compostos com propriedades de fluorescência diferentes. Após a hibridização, cada cromossomo fluoresce em uma cor diferente.

pirimidina Um tipo de base nitrogenada; as bases pirimidínicas no DNA são citosina e timina.

placa (1) Um frasco achatado usado para o cultivo de micróbios. (2) Para espalhar células em uma superfície de meio sólido em uma placa. (3) Uma área clara em uma camada de bactérias, deixada por lise da bactéria por meio de infecções progressivas por um fago e seus descendentes.

planta dióica Uma espécie de planta na qual os órgãos masculinos e femininos estão em plantas separadas.

plaqueamento Espalhar células de microrganismos (bactérias, fungos) em um frasco de meio nutritivo de modo a permitir que cada célula forme uma colônia visível.

plaqueamento réplica Em genética microbiana, um modo de triar colônias dispostas em uma placa master para ver se são mutantes em outros ambientes; é usada uma almofada de feltro para transferir as colônias para novas placas.

plasmídeo Um DNA extracromossômico de replicação autônoma.

plasmídeo F' *Veja fator F'.*

plasmídeo R Um plasmídeo contendo um dentre vários transposons que possuem genes de resistência.

plasmídeo Ti Um plasmídeo circular de *Agrobacterium tumefaciens* que permite que a bactéria infecte células vegetais e produza um tumor (galha).

ploidia O número de conjuntos cromossômicos.

poky Um mutante mitocondrial de crescimento lento em *Neurospora*.

poliacrilamida Um material usado para fazer géis de eletroforese para a separação de misturas de macromoléculas.

polidactilia Mais de cinco dedos e/ou artelhos. Herdada como um fenótipo autossômico dominante.

poligene (locus de característica quantitativa) Um gene que apresenta alguns alelos capazes de interagir aditivamente com alelos em outros loci para afetar um fenótipo (característica) mostrando distribuição contínua.

poligenes *Veja hipótese de fator múltiplo.*

poliligador Uma sequência de DNA vetor contendo vários sítios únicos de corte com enzima de restrição que é conveniente para inserir DNA exógeno. Às vezes chamado também de MCS (sítio de clonagem múltipla).

polimerase de bypass Uma DNA polimerase que continua a replicar o DNA além do sítio de dano que pararia a replicação pela polimerase replicativa normal. As polimerases de *bypass* contribuem para um mecanismo de tolerância ao dano chamado síntese de DNA translesão.

polimerase distributiva Uma polimerase que pode adicionar apenas alguns nucleotídeos antes de se dissociar do molde.

polimerases translesão Uma família de DNA polimerases que podem continuar a replicar o DNA além de um sítio de dano que pararia a replicação pela polimerase replicativa normal. Também conhecidas com *polimerases de bypass*.

polimorfismo A ocorrência em uma população (ou entre populações) de várias formas fenotípicas associadas a alelos de um gene ou homólogos de um cromossomo.

polimorfismo de comprimento de fragmento de restrição (RFLP) Uma diferença em sequência de DNA entre indivíduos ou haplótipos que é reconhecida como tamanhos diferentes de fragmentos de restrição. Por exemplo, uma substituição de um par de nucleotídeos pode causar um sítio de reconhecimento de enzima de restrição presente em um alelo de um gene e ausente no outro. Conseqüentemente, uma sonda para essa região do DNA hibridizará com fragmentos de tamanhos diferentes dentro das digestões de restrição de DNA desses dois alelos.

polimorfismo de comprimento de sequência curta (SSLP) A presença de números diferentes de elementos curtos ou simplesmente repetitivos (DNA mini e microsatélite) em um locus particular em cromossomos homólogos diferentes; os heterozigotos servem como marcadores úteis para o mapeamento do genoma.

polimorfismo de comprimento de sequência simples (SSLP) A existência na população de indivíduos apresentando números diferentes de cópias de uma sequência simples, curta, em um locus cromossômico.

polimorfismo de DNA Uma variação de ocorrência natural na sequência de DNA em determinado local no genoma.

polimorfismo de nucleotídeo único (SNP) Uma diferença de um par de nucleotídeos em determinado locus nos genomas de dois ou mais indivíduos de ocorrência natural.

polimorfismo genético Diferenças genéticas de ocorrência natural entre membros individuais de uma população.

polipeptídeo Uma cadeia de aminoácidos ligados; uma proteína.

poliplóide Uma célula tendo três ou mais conjuntos cromossômicos ou um organismo composto de tais células.

polissacarídeo Um polímero biológico composto de subunidades de açúcar – por exemplo, amido ou celulose.

ponta amino A ponta de uma proteína tendo um grupo amino livre. Uma proteína é sintetizada a partir da ponta amino na extremidade 5' de uma molécula de mRNA até a ponta carboxila perto da ponta 3' do mRNA durante a tradução.

ponta carboxila O final de uma proteína tendo um grupo carboxila livre. A ponta carboxila é codificada pela ponta 3' do mRNA e é a última parte da proteína a ser sintetizada na tradução.

ponte anafásica Em um cromossomo dicêntrico, o segmento entre os centrômeros que estão sendo levados para pólos opostos na divisão nuclear.

ponte de hidrogênio Uma ligação fraca na qual um átomo compartilha um elétron com um átomo de hidrogênio; as pontes de hidrogênio são importantes na especificidade do pareamento de bases nos ácidos nucleicos e na determinação da forma da proteína.

ponte dicêntrica Em um cromossomo dicêntrico, o segmento entre os centrômeros que está sendo levado para pólos opostos na divisão nuclear.

ponto quente Uma parte de um gene que mostra uma tendência muito alta em tornar-se um sítio mutante, seja espontaneamente ou sob a ação de um mutágeno em particular.

população Um grupo de indivíduos que se reproduzem uns com os outros para produzir a geração seguinte.

populações alopátricas Populações que são separadas no espaço e que não trocam genes.

portador Um organismo que possui um alelo mutante mas não o expressa em seu fenótipo, devido a um alelo dominante; assim, um indivíduo com genótipo *A/a* é um portador de *a* se houver dominância completa de *A* sobre *a*.

posterior Um modo de descrever a localização relativa de um sítio em um DNA ou molécula de RNA. Um sítio posterior está situado mais perto da ponta 3' de uma unidade de transcrição.

pré-mRNA *Veja transcrito primário.*

prendedor deslizante Uma proteína acessória de replicação em bactérias que envolve o DNA como uma rosquinha.

primase Uma enzima que faz primers de RNA na replicação do DNA.

primeira geração filial (F₁) Os indivíduos da prole que surgem de um cruzamento de duas linhagens diplóides homozigotas.

Primeira Lei de Mendel Os dois membros de um par de genes segregam-se um do outro na meiose; cada gameta tem uma probabilidade igual de obter qualquer um dos membros do par de genes.

primer Um oligonucleotídeo de RNA ou DNA que pode servir como um molde para a síntese de DNA pela DNA primase quando helicoidizado a uma molécula maior de DNA.

primossomo Um complexo proteico na forquilha de replicação cujo componente central é a *primase*.

procarionte Um organismo composto de uma célula procariótica, tal como uma bactéria ou alga cianofíceia.

processamento co-transcricional A transcrição simultânea e processamento do pré-mRNA eucariótico.

processamento de RNA O termo coletivo para as modificações de RNA eucariótico, incluindo a adição do cap e a recomposição, que são necessárias antes que o RNA possa ser transportado para o citoplasma para a tradução.

processamento pós-transcricional Modificações de grupos laterais de aminoácidos após uma proteína ter sido liberada do ribossomo.

produto da meiose Uma das células (geralmente quatro) formadas por duas divisões meióticas.

produtos de crossing As células produto de meiose com cromossomos que participaram de um crossing.

profago Um “cromossomo” de fago inserido como parte da estrutura linear do DNA cromossômico de uma bactéria.

prófase O primeiro estágio da divisão nuclear durante o qual os cromossomos se condensam e tornam-se visíveis.

proflavina Um mutágeno que tende a produzir mutações de mudança de matriz de leitura.

projeto de drogas baseado na estrutura O uso de informações básicas sobre processos celulares e a maquinaria para desenvolver drogas.

Projeto Genoma Um esforço em larga escala, geralmente multilaboratorial necessário para seqüenciar um genoma complexo.

promotor Uma região reguladora a curta distância da ponta 5' de um gene e que atua como sítio de ligação para a RNA polimerase.

proporção mendeliana Uma proporção de fenótipos da prole que corresponde ao funcionamento das leis de Mendel.

proporção não-mendeliana Uma proporção incomum de fenótipos da prole que não se adapta à simples operação das leis de Mendel; por exemplo, mutante: proporções tipo selvagem 3:5, 5:3, 6:2 ou 2:6 em tétrades indicam que ocorreu conversão gênica.

propósito Em um heredograma, a pessoa que primeiro chamou a atenção do geneticista.

propriedade Uma característica de um organismo, tal como tamanho, cor, forma ou atividade de enzima.

proteína Uma macromolécula composta de uma ou mais cadeias de aminoácidos; o principal componente da expressão fenotípica.

proteína-1 de heterocromatina (HP-1) Uma proteína necessária para a manutenção da heterocromatina.

proteína acessória Uma proteína associada a DNA polimerase III de *E. coli* que não é parte do cerne catalítico.

proteína ativadora do catabolismo Uma proteína que se liga ao cAMP em concentrações baixas de glicose e liga-se ao promotor *lac* para facilitar a ação da RNA polimerase.

proteína de ligação a TATA (TBP) Um fator geral de transcrição que se liga ao TATA box e ajuda a atrair outros fatores gerais de transcrição e RNA polimerase II para promotores eucarióticos.

proteína de ligação unifilamentar (SSB) Uma proteína que se liga a DNA unifilamentares e impede que dupla hélice seja reconstituída antes da replicação.

proteína estrutural Uma proteína que funciona na estrutura de organismos celulares.

proteína fibrosa Uma proteína com uma forma linear tal como os componentes dos cabelos e músculos.

proteína globular Uma proteína com uma estrutura compacta, tal como uma enzima ou anticorpo.

proteína hélice-alça-hélice Veja proteínas HLH.

proteínas HLH (hélice-alça-hélice) Uma família de proteínas na qual uma parte do polipeptídeo forma duas hélices separadas por uma alça (o domínio HLH); essa estrutura atua como um domínio de ligação ao DNA específico de seqüência. As proteínas HLH são tidas como agindo como fatores de transcrição.

proteoma O conjunto completo de genes codificantes de proteína em um genoma.

proteômica O estudo sistemático do proteoma.

proto-oncogene A contraparte celular normal de um gene que pode ser mutado para tornar-se um oncogene dominante.

protoplasto Uma célula vegetal da qual foi removida a parede.

prototrófica Uma linhagem de organismos que irão proliferar em meio mínimo (*compare com auxotrófica*).

provírus O genoma de DNA cromossomicamente inserido de um retrovírus.

pseudodominância O súbito aparecimento de um fenótipo recessivo em um pedigree, devido a uma deleção de um gene dominante mascarador.

pseudogene Um gene mutacionalmente inativo para o qual não existe contraparte funcional nas populações tipo selvagem.

pseudogene processado Um pseudogene que surge pela transcrição reversa de um mRNA maduro e sua integração ao genoma.

pseudoligação O surgimento de ligação de dois genes em cromossomos translocados.

punção de vilosidades coriônicas (CVS) Um procedimento que visa obter amostra de placenta a fim de obter tecido fetal para análise de DNA de tecido fetal para análise cromossômica no diagnóstico pré-natal de distúrbios genéticos.

purina Um tipo de base nitrogenada; as bases purínicas no DNA são adenina e guanina.

QTL Veja locus de característica quantitativa.

quadrado de Punnett Uma grade usada como representação gráfica dos zigotos da prole resultantes das fusões de gametas diferentes em um cruzamento específico.

quebra bifilamentar Uma quebra de DNA que corta os arcabouços açúcar-fosfato de ambos os filamentos da dupla hélice de DNA.

quiasma Uma estrutura em forma de cruz comumente observada entre cromátides não-irmãs na meiose; o sítio do crossing over.

quilobase (kb) Mil pares de nucleotídeos.

quimera Um tecido contendo dois ou mais tipos celulares genéticos ou um organismo individual composto de tais tecidos.

reação em cadeia da polimerase (PCR) Um método *in vitro* para amplificar um segmento específico de DNA que usa dois primers que se hibridizam com pontas opostas do segmento em polaridade oposta e, durante ciclos sucessivos, iniciam a replicação exponencial apenas desse segmento.

rearranjo A produção de cromossomos anormais pela quebra e reunião incorreta dos segmentos cromossômicos; são exemplos as inversões, deleções e translocações.

rearranjo cromossômico Uma mutação cromossômica na qual partes do cromossomo estão em nova justaposição.

rearranjo desbalanceado Um rearranjo no qual material cromossômico é ganho ou perdido em um conjunto cromossômico.

rearranjos balanceados Aqueles nos quais não há perda nem ganho de material genético.

receptor Veja interação ligando-receptor.

receptor transmembranar Uma proteína inserida na membrana citoplasmática de uma célula, com a parte extracelular da proteína tendo a habilidade de se ligar a um ligando e a parte intracelular tendo uma atividade (tal como a cinase proteica) que pode ser induzida na união do ligando.

receptora A bactéria que recebe DNA em uma transferência unilateral entre as células; os exemplos são F⁻ em uma conjugação ou a célula transduzida em uma transdução mediada por fago.

recombinação (1) Em geral, qualquer processo em uma célula diplóide ou parcialmente diplóide que gera um novo gene ou combinações cromossômicas não encontradas previamente nessa célula ou em seus genitores. (2) Na meiose, o processo que gera um produto haplóide na meiose cujo genótipo é diferente dos dois genótipos haplóides que constituem o diplóide meiótico.

recombinação de fago A produção de genótipos de fagos recombinantes como resultado de uma dupla infecção de uma bactéria com genótipos de fago “parentais” diferentes.

recombinação homóloga não-alélica (NAHR) Crossing over entre curtas unidades homólogas encontradas em diferentes loci cromossômicos.

recombinação meiótica Recombinação por distribuição ou crossing over na meiose.

recombinante Refere-se a um organismo individual ou célula tendo um genótipo produzido por recombinação.

recomposição (splicing) Uma reação que remove íntrons e une éxons no RNA.

recomposição alternativa Um processo pelo qual diferentes RNA mensageiros são produzidos pelo mesmo transcrito primário, por meio de variações no padrão de recomposição do transcrito. Múltiplas “isoformas” de mRNA podem ser produzidas em uma única célula, ou isoformas diferentes podem apresentar diferentes padrões histoespecíficos de expressão. Se éxons alternativos caem dentro de matrizes de leitura abertas de isoformas de mRNA, serão produzidas proteínas diferentes por mRNA alternativos.

recomposição de RNA Uma reação encontrada amplamente em eucariontes que removem íntrons e juntam os éxons no RNA.

“recortar e colar” (“cut and paste”) Termo descritivo para um mecanismo de transposição no qual o transposon classe 2 (DNA) é removido (*cut*) do sítio doador e inserido (*pasted*) em um novo tecido-alvo.

recuperação de mutante *Veja complementação funcional.*

reelicoïdização Realinhamento espontâneo de dois filamentos de DNA para reconstituir uma dupla hélice que foi desnaturada.

refinamento do destino O processo pelo qual as decisões são ajustadas de modo que todos os tipos celulares necessários sejam destinados aos locais espaciais apropriados e número de células.

região não-traduzida 3’ (3’UTR) A região do RNA transcrito na ponta 3’ posterior ao sítio de término da tradução.

região não-traduzida 5’ (5’UTR) A região do RNA transcrito na ponta 5’ antecedente ao sítio de início da tradução.

região pseudo-autossômica Uma região de um cromossomo sexual que é homóloga a uma região no outro cromossomo sexual heterogamético; elas atuam como substratos do pareamento meiótico e crossing over.

região reguladora Sequência não-codificante de um gene à qual se ligam várias proteínas que causam a transcrição do gene no tempo e lugar corretos.

regra da soma A probabilidade de que um ou outro de dois eventos mutuamente exclusivos ocorra é a soma de suas probabilidades individuais.

regra do produto A probabilidade de que dois eventos que ocorrem simultaneamente sejam o produto de probabilidades individuais.

regra GU-AG Assim chamada porque os dinucleotídeos GU e AG estão quase sempre nas pontas 5’ e 3’, respectivamente, de íntrons, onde eles são reconhecidos por componentes do spliceossomo.

relação Em estatística, a correlação entre duas quantidades mensuráveis.

relógio molecular A taxa constante de substituição de aminoácidos nas proteínas ou nucleotídeos nos ácidos nucleicos em um longo tempo evolucionário.

remodelagem de cromatina Mudanças na posição do nucleossomo ao longo do DNA.

reparo de excisão O reparo de uma lesão do DNA pela remoção do segmento de DNA ausente e sua substituição por um segmento tipo selvagem.

reparo de excisão de nucleotídeo acoplado a transcrição (TC-NER) Uma forma de reparo de excisão de nucleotídeo que é ativada por complexos de transcrição parados e corrige o dano ao DNA em regiões transcritas do genoma.

reparo genômico global (GGR) Um tipo de reparo por excisão de nucleotídeo que ocorre em sequências não-transcritas.

reparo por excisão de base Uma das várias vias de reparo por excisão. Nessa via, pequenas distorções de pares de bases são reparadas pela criação de sítios apurínicos seguidos de síntese de reparo.

repetição de trinucleotídeo *Veja expansão de trinca.*

replicação Síntese de DNA.

replicação conservativa Um modelo não-comprovado de síntese de DNA sugerindo que metade das moléculas de DNA deveriam ter ambos os filamentos compostos de nucleotídeos recém-polimerizados.

replicação dispersiva Um modelo não-comprovado de síntese de DNA sugerindo a intercalação mais ou menos aleatória de segmentos parentais e novos nas moléculas-filhas de DNA.

replicação em círculo rolante Um modo de replicação usado por algumas moléculas de DNA circulares em bactérias (tais como plasmídeos) nas quais o círculo parece girar à medida que sai de um filamento leading contínuo.

replicação semiconservativa O modelo estabelecido de replicação do DNA no qual cada molécula bifilamentar é composta de um filamento parental e um filamento recém-polimerizado.

replissomo A máquina molecular na forquilha de replicação que coordena as várias reações necessárias para a replicação rápida e precisa do DNA.

representação aleatória Representação entre indivíduos nos quais a escolha de um parceiro não é influenciada pelos genótipos (com relação a genes específicos em estudo).

repressão catabólica A inativação de um óperon causada pela presença de grandes quantidades do produto final catabólico do óperon.

repressor Uma proteína que se liga a um elemento cis-trans tal como um operador ou um silenciador, impedindo assim a transcrição de um promotor adjacente.

reprodução interrompida Uma técnica usada para mapear genes bacterianos determinando a sequência na qual os genes doadores entram nas células receptoras.

reprodução preferencial negativa Reprodução preferencial entre parceiros fenotipicamente diferentes.

reprodução preferencial positiva Uma situação na qual fenótipos iguais se reproduzem mais freqüentemente do que o esperado pelo acaso.

resposta de seleção A quantidade de mudança no valor médio de alguma característica fenotípica entre a geração parental e a geração prole como resultado da seleção dos genitores.

retroelemento O nome geral para os elementos de transposição classe 1 que se movem através de um RNA intermediário.

retrotransposição Um mecanismo de transposição caracterizado pelo fluxo reverso de informação do RNA para o DNA.

retrotransposon Um elemento de transposição que usa transcriptase reversa para transpor através de um RNA intermediário.

retrotransposon LTR Um tipo de elemento de transposição classe 1 que termina em longas repetições terminais e codifica várias proteínas, incluindo transcriptase reversa.

retrovírus Um vírus com RNA que se replica primeiro, sendo convertido em dupla hélice de DNA.

reversão A produção de um gene tipo selvagem a partir de um gene mutante.

revertente Um alelo com função tipo selvagem que surge de mutação de um alelo mutante; causado pela reversão completa do evento original ou por uma segunda mutação compensatória.

RF *Veja frequência de recombinantes; fator de liberação.*

RFLP *Veja polimorfismo de comprimento de fragmento de restrição.*

ribose O açúcar pentose do RNA.

ribossomo Uma organela complexa que catalisa a tradução do RNA mensageiro em uma sequência de aminoácidos; composta de proteínas mais rRNA.

ribozima Um RNA com atividade enzimática – por exemplo, as moléculas de RNA auto-recompostas em *Tetrahymena*.

RISC (complexo de silenciamento induzido por RNA) Um complexo de proteína multissubunitária que se associa a siRNA e é guiado para um mRNA-alvo por complementaridade de base. O mRNA-alvo é clivado pela atividade RISC.

RNA (ácido ribonucleico) Um ácido nucleico unifilamentar similar ao DNA, mas tendo o açúcar ribose, em vez de desoxirribose, e uracil, em vez de timina, como uma de suas bases.

RNA de interferência (RNAi). Um modo de avaliar o funcionamento de um gene introduzindo construções transgênicas especiais para inativar seu mRNA.

RNA final mRNA eucariótico que é totalmente processado e pronto para ser exportado do núcleo e traduzido no citoplasma.

RNA funcional Um tipo de RNA que tem um papel sem ser traduzido.

RNA mensageiro *Veja mRNA.*

RNA polimerase Uma enzima que catalisa a síntese de um filamento de RNA a partir de um molde de DNA. Os eucariontes possuem várias classes de RNA polimerase; genes estruturais codificando proteínas são transcritos pela RNA polimerase II.

RNA polimerase holoenzima O complexo multissubunitário bacteriano composto de quatro subunidades do cerne da enzima mais o fator σ .

RNA ribossômico Veja **rRNA**.

RNA transportador Veja **tRNA**.

RNAi Veja **RNA de interferência**.

robótica A aplicação da tecnologia automatizada para coletar dados biológicos em larga escala, tais como o Projeto Genoma.

rRNA (RNA ribossômico) Uma classe de moléculas de RNA, codificadas no organizador nucleolar, que tem um papel integral (mas pouco compreendido) na estrutura e função do ribossomo.

ruído desenvolvimental Variação no resultado do desenvolvimento como consequência de eventos aleatórios na divisão celular, movimento celular e pequenas diferenças no número e localização de moléculas dentro das células.

segmentação O processo pelo qual o número correto e os tipos de segmentos são estabelecidos em um animal superior em desenvolvimento.

segmento Em animais superiores, tais como anelídeos, artrópodos, cordados, uma das unidades repetidas ao longo do eixo anterior-posterior do corpo.

segregação (1) Citologicamente, a separação de estruturas homólogas. (2) Geneticamente, a produção de dois fenótipos separados, correspondendo a dois alelos de um gene, seja em indivíduos diferentes (segregação meiótica) ou em tecidos diferentes (segregação mitótica).

segregação adjacente-1 Em uma translocação recíproca, a passagem de um cromossomo translocado e um normal para cada um dos pólos.

segregação alternada Em uma translocação recíproca, a passagem de ambos os cromossomos normais para um pólo e de ambos os cromossomos translocados para o outro pólo.

segregação citoplasmática Segregação na qual células-filhas geneticamente diferentes surgem de um genitor que é um citohet.

segregação igual Números iguais de genótipos da prole atribuíveis à separação de dois alelos de um gene na meiose.

segunda geração filial (F_2) A prole de um cruzamento entre dois indivíduos da geração F_1 .

Segunda Lei de Mendel A lei da distribuição independente; pares de genes não-ligados ou distantemente ligados separam-se independentemente na meiose.

seleção (1) Um procedimento experimental no qual apenas um tipo específico de mutante pode sobreviver. (2) A produção de números médios diferentes de prole por genótipos diferentes em uma população como resultado de propriedades fenotípicas diferentes desses genótipos.

seleção artificial Cruzamento de gerações sucessivas por deliberada seleção humana de alguns fenótipos ou genótipos como genitores de cada geração.

seleção balanceadora Seleção natural que resulta em um equilíbrio intermediário estável de frequências alélicas.

seleção diferencial A diferença entre a média de uma população e a média dos membros individuais selecionados para serem genitores da geração seguinte.

seleção direcional Seleção que muda a frequência de um alelo em uma direção constante, aproximando-se ou afastando-se da fixação desse alelo.

seleção familiar Uma técnica de cruzamento para selecionar um par com base no desempenho médio de sua prole.

seleção genética Veja **sistema seletivo**.

seleção natural A taxa diferencial de reprodução de tipos diferentes em uma população como resultado de diferentes características fisiológicas, anatômicas ou comportamentais dos tipos.

seleção negativa A eliminação de uma característica deletéria de uma população por seleção natural.

seleção purificadora Seleção natural que remove variantes deletérias de um DNA ou sequência de proteínas, reduzindo assim a diversidade genética.

seleção truncada Uma técnica de cruzamento pela qual organismos individuais, nos quais a expressão quantitativa de um fenótipo está acima ou abaixo de um certo valor (o ponto de truncação), são selecionados para a geração seguinte.

semi-esterilidade O fenótipo de um organismo heterozigoto para alguns tipos de *anomia cromossômica*; expresso como um número reduzido de gametas viáveis e, portanto, de reduzida fertilidade.

sequência contig Um grupo de segmentos clonados superpostos.

sequência de ativação antecedente (UAS) Uma sequência de DNA de levedura situada a 5' do gene promotor; um fator de transcrição liga-se à UAS para regular positivamente a expressão gênica.

sequência de consenso A sequência de nucleotídeos de um segmento de DNA que está de acordo com a mesma sequência de leitura do mesmo segmento de pessoas diferentes.

sequência de localização nuclear (NLS) Parte de uma proteína necessária para seu transporte do citoplasma para o núcleo.

sequência de repetição invertida (IR) Uma sequência encontrada em forma idêntica (mas invertida) – por exemplo, as pontas opostas de um transposon de DNA.

sequência de sinal A sequência aminoterminal de uma proteína secretada; é necessária para o transporte da proteína através da membrana celular.

sequência expressa marcada (EST) Um clone de cDNA para o qual apenas as pontas 5' e/ou 3' foram sequenciadas; usada para identificar as pontas dos transcritos na análise genômica.

sequência intercalar Um íntron; um segmento de função amplamente desconhecida dentro de um gene. Esse segmento é inicialmente transcrito, mas o transcrito não é encontrado no mRNA funcional.

sequência IR Veja **sequência de repetição invertida**.

sequência líder A sequência na ponta 5' de um mRNA que não é traduzida em proteína.

sequência Shine-Dalgarno Uma sequência curta no RNA bacteriano que precede o códon de iniciação AUG e serve para posicionar corretamente esse códon no sítio P do ribossomo por pareamento (pela complementaridade das bases) com a ponta 3' do RNA 16S na subunidade 30S do ribossomo.

sequenciamento de clone ordenado O sequenciamento de um conjunto de clones, tais como uma via mínima, que corresponde a um segmento genômico, um cromossomo ou um genoma.

sequenciamento de Sanger Veja **sequenciamento dideoxi**.

sequenciamento dideoxi O método mais popular de sequenciamento de DNA. Ele usa trifosfatos de dideoxinucleotídeo misturados com trifosfatos padrão de nucleotídeos para produzir uma escada de filamentos de DNA cuja síntese é bloqueada em tamanhos diferentes. Esse método foi incorporado a máquinas automatizadas de síntese de DNA. Também chamado de sequenciamento Sanger em homenagem a seu inventor, Frederick Sanger.

sequenciamento genômico Um projeto em grande escala para sequenciar um genoma completo.

sequenciamento shotgun de genoma inteiro (WGS) O sequenciamento das pontas de clones sem relação a nenhuma informação sobre o local dos clones.

séries alélicas O conjunto de alelos conhecidos de um gene.

setor Uma área de tecido cujo fenótipo é detectavelmente diferente do fenótipo do tecido vizinho.

sexo heterogamético O sexo que tem cromossomos sexuais heteromórficos (p. ex., XY) e assim produz dois tipos diferentes de gametas com relação a cromossomos sexuais.

sexo homogamético O sexo com cromossomos sexuais homólogos (p. ex., XX).

silenciamento de transgene Refere-se à presença de um gene exógeno em um organismo transgênico que não produz um mRNA ou produto proteico devido a modificações epigenéticas.

silenciamento gênico Um gene que não é expresso devido à regulação epigenética. Ao contrário dos genes que são mutantes devido a alterações na sequência de DNA, os genes inativados por silenciamento podem ser reativados.

sinapse Pareamento próximo de homólogos na meiose.

sincício Uma única célula tendo vários núcleos.

síndrome de Cockayne Um distúrbio genético causado por defeitos no sistema de reparo de excisão de nucleotídeo levando a sintomas de envelhecimento prematuro. As pessoas com síndrome de Cockayne têm uma mutação em uma de duas proteínas tidas como reconhecendo complexos de transcrição que estão bloqueados devido a dano ao DNA.

síndrome de Down Um fenótipo humano anormal, que inclui retardo mental, devido a uma trissomia do cromossomo 21; mais comum em crianças nascidas de mães com mais idade.

síndrome de Klinefelter Um fenótipo masculino humano anormal devido a um cromossomo X extra (XXY).

síndrome de Turner Um fenótipo feminino humano anormal produzido pela presença apenas de um cromossomo X.

SINE *Veja elemento curto intercalado.*

sintenia Uma situação na qual os genes estão dispostos em blocos similares em espécies diferentes.

síntese de DNA translesão Um mecanismo de tolerância ao dano nos eucariontes que usa *polimerases de bypass* para replicar o DNA além de um sítio de dano.

siRNA *Veja pequeno RNA de interferência.*

sistema diíbrido Um par de vetores de *Saccharomyces cerevisiae* (levedura) usado para detectar interações proteína-proteína. Cada vetor leva o gene para uma proteína exógena diferente em teste; se os vetores se unirem fisicamente, é transcrito um gene repórter.

sistema de reparo de mal pareamento Um sistema para reparar o dano ao DNA que já foi replicado.

sistema de reparo dependente de homologia Um mecanismo de reparo de DNA que depende da complementaridade ou homologia do filamento-molde com o filamento que está sendo reparado.

sistema de reparo por excisão de nucleotídeo Uma via de reparo de excisão que quebra as ligações fosfodiéster em ambos os lados de uma base danificada, com remoção dessa base e de várias, de ambos os lados, seguida de replicação de reparo.

sistema diíbrido de levedura *Veja sistema diíbrido.*

sistema seletivo Uma técnica de seleção mutacional que enriquece a frequência de genótipos específicos (geralmente raros) estabelecendo condições ambientais que impedem o crescimento ou sobrevivência de outros genótipos.

sistema SOS (reparo) Um processo propenso a erro no qual uma polimerase de *bypass* replica um dano ao DNA em uma forquilha de replicação bloqueada pela inserção de bases inespecíficas.

sítio A *Veja sítio de ligação de aminoacil-tRNA.*

sítio alostérico Um sítio em uma proteína ao qual se liga uma pequena molécula, causando uma mudança na conformação da proteína que modifica a atividade de seu sítio ativo.

sítio apirimidínico Um sítio de DNA que perdeu uma pirimidina.

sítio apurínico Um sítio de DNA que perdeu uma purina.

sítio ativo A parte de uma proteína que deve ser mantida em uma forma específica para que a proteína seja funcional; por exemplo, em uma enzima, a parte à qual se liga o substrato.

sítio de clonagem múltipla *Veja poliligador.*

sítio de ligação Região à qual o profago se integra.

sítio de ligação de aminoacil-tRNA Um sítio no ribossomo que se liga ao aminoacil-tRNA que chega. O anticódon de cada aminoacil-tRNA que chega parecia com o códon do mRNA. Também chamado de sítio A.

sítio de saída (E) O sítio do ribossomo onde o tRNA desacilado pode ser encontrado.

sítio E *Veja sítio de saída.*

sítio mutante A área danificada ou alterada dentro de um gene mutante.

sítio P *Veja sítio peptidil.*

sítio peptidil (P) O sítio no ribossomo ao qual se liga um tRNA em uma cadeia polipeptídica crescente.

SNP *Veja polimorfismo de nucleotídeo único.*

snRNA *Veja pequeno RNA nuclear.*

sobredominância Uma relação fenotípica na qual a expressão fenotípica do heterozigoto é maior que ambos os homozigotos.

sonda Segmento de ácido nucleico marcado que pode ser usado para identificar moléculas específicas de DNA portando a sequência complementar, geralmente por auto-radiografia ou fluorescência.

sondagem O método mais amplamente usado para detectar macromoléculas específicas: uma mistura de macromoléculas (DNA, RNA ou proteína) é exposta a uma molécula, a sonda, que se ligará apenas à macromolécula procurada.

spliceossomo O complexo processo ribonucleoproteico que remove íntrons de mRNA eucarióticos.

SSB *Veja proteína de ligação unifilamentar.*

SSLP *Veja polimorfismo de comprimento de sequência curta.*

subdominância Uma relação fenotípica na qual a expressão fenotípica do heterozigoto é menor do que a de ambos os homozigotos.

substituição conservativa Substituição de um par de nucleotídeos dentro de uma região codificante de proteína que leva à substituição de um aminoácido por outro de propriedades químicas similares.

substituição de par de bases *Veja substituição de par de nucleotídeos.*

substituição de par de nucleotídeos A substituição de um par de nucleotídeos específico por um par diferente; em geral é mutagênica.

substituição gênica A inserção de um transgene geneticamente modificado no lugar de um gene residente; geralmente obtida por um crossing duplo.

substituição não-conservativa Substituição de par de nucleotídeos dentro de uma região codificante de proteína que leva à substituição de um aminoácido por outro com propriedades químicas diferentes.

substituição não-sinônima Substituição mutacional de um aminoácido por outro com propriedades químicas diferentes.

substituição sinônima *Veja mutação sinônima.*

subunidade Como usada no Capítulo 9, um único polipeptídeo em uma proteína contendo vários polipeptídeos.

sulco maior O maior dos dois sulcos na dupla hélice de DNA.

sulco menor O menor dos dois sulcos na dupla hélice de DNA.

supercontig *Veja arcabouço (2).*

superfície adaptativa *Veja panorama adaptativo.*

super-hélice Uma molécula bifilamentar fechada de DNA que é girada sobre si mesma.

supressão A produção de um fenótipo próximo do tipo selvagem pela adição de outra mutação a um genótipo fenotipicamente anormal.

supressor Uma mutação secundária que pode cancelar o efeito de uma mutação primária, resultando em um fenótipo tipo selvagem.

supressora sem sentido Uma mutação que produz um tRNA alterado que irá inserir um aminoácido na tradução em resposta a um códon sem sentido.

T (1) *Veja timidina; timina.* (2) *Veja tetratipo.*

TATA boxe Uma sequência de DNA encontrada em muitos genes eucarióticos que está situada a cerca de 30 pb do sítio de início da transcrição.

tautômeros Isômeros de uma molécula tal como as bases do DNA que diferem nas posições dos átomos e ligações entre os átomos.

taxa de mutação O número de eventos mutacionais por cópia de gene em uma população por unidade de tempo (p. ex., por geração celular).

TBP *Veja proteína de ligação a TATA.*

TC-NER *Veja reparo de excisão de nucleotídeo acoplado a transcrição.*

T-DNA Uma parte do plasmídeo Ti que é inserida no genoma da célula da planta hospedeira.

técnica shotgun A clonagem de um grande número de fragmentos diferentes de DNA com um prelúdio para selecionar um tipo particular de clone para estudo intensivo.

tecnologia do DNA As técnicas coletivas para obter, amplificar e manipular fragmentos específicos de DNA.

telófase O último estágio da divisão nuclear, quando são reconstituídos os núcleos-filhos.

telomerase Uma enzima que, com o uso de um pequeno RNA especial como molde, adiciona unidades repetitivas às pontas de cromossomos lineares para evitar o encurtamento após a replicação.

telômero A ponta ou final de um cromossomo.

temperatura permissiva A temperatura na qual um alelo mutante sensível à temperatura é expresso do mesmo modo que o alelo tipo selvagem.

temperatura restritiva A temperatura na qual uma mutação sensível à temperatura expressa o fenótipo mutante.

tendência central O valor de uma variável contínua ao redor do qual se agrupam os valores individuais de distribuição.

teoria cromossômica da herança Uma teoria unificadora que diz que os padrões de herança podem geralmente ser explicados supondo que os genes estão situados em sítios específicos nos cromossomos.

terapia gênica A correção de uma deficiência genética em uma célula pela adição de um novo DNA e sua inserção no genoma. Técnicas diferentes têm o potencial de fazer terapia gênica só nos tecidos somáticos ou, alternativamente, para corrigir a deficiência genética no zigoto, corrigindo assim também a linhagem germinativa.

terapia gênica de linhagem germinativa *Veja terapia gênica.*

terapia gênica somática *Veja terapia gênica.*

término (1) O final representado pelo último monômero adicionado na síntese unidirecional de um polímero tal como RNA ou um polipeptídeo. (2) O último estágio da transcrição; resulta na liberação do RNA e da RNA polimerase de um molde de DNA.

testador Um organismo individual homozigoto para um ou mais alelos recessivos; usado em um cruzamento-teste.

teste de Ames Um modo para testar se um composto químico é mutagênico expondo linhagens bacterianas mutantes especiais ao produto formado pela digestão desse composto por extrato de fígado e, então, contando o número de colônias. Apenas novas mutações, supostamente produzidas pelo composto, podem produzir revertentes ao tipo selvagem capaz de formar colônias.

teste de complementação Um teste para determinar se duas mutações estão em genes diferentes (elas se complementam) ou no mesmo gene (elas não se complementam).

teste de flutuação Um teste usado em micróbios para estabelecer a natureza aleatória da mutação ou para medir as taxas de mutação.

teste diíbrido Um método para detectar interações proteína-proteína, tipicamente feito em leveduras.

teste do qui-quadrado (χ^2) Um teste estatístico usado para determinar a probabilidade de obter as proporções observadas por acaso, sob uma hipótese específica.

tétrade (1) Quatro cromátides homólogas em um feixe na primeira prófase meiótica e metáfase. (2) Os quatro produtos celulares haplóides de uma única meiose.

tétrade linear Uma tétrade que resulta de divisões nucleares meióticas e pós-meióticas de tal modo que os produtos-irmãos permanecem adjacentes uns aos outros (sem passagem de núcleos).

tétrade não-linear Uma tétrade na qual os produtos meióticos não estão em determinada ordem.

tetrâmero Uma proteína que consiste em quatro subunidades polipeptídicas.

tetraplóide Uma célula tendo quatro conjuntos cromossômicos; um organismo composto de tais células.

tetratipo (T) Um tipo de tétrade contendo quatro genótipos diferentes: dois parentais e dois recombinantes.

timidina (T) Um nucleosídeo tendo timina como sua base.

timina (T) Uma base pirimidínica que faz par com adenina.

tipo selvagem O genótipo ou fenótipo que é encontrado na natureza ou no estoque padrão do laboratório para determinado organismo.

tipos reprodutivos O equivalente em organismos inferiores aos sexos nos organismos superiores; os tipos reprodutivos diferem tipicamente apenas fisiologicamente, não na forma física.

Tn *Veja transposon.*

topoisomerase Uma enzima que pode cortar e reconstituir arcabouços polinucleotídicos no DNA para permitir que assumam uma configuração mais relaxada.

tradução A produção mediada por ribossomo e tRNA de um polipeptídeo cuja sequência de aminoácidos é derivada da sequência códon de uma molécula de mRNA.

transcrição A síntese de RNA a partir de um molde de DNA.

transcriptase reversa Uma enzima que catalisa a síntese de um filamento de DNA a partir de um molde de RNA.

transcrito A molécula de RNA copiada do filamento-molde de DNA pela RNA polimerase.

transcrito primário (pré-mRNA) RNA eucariótico antes de ser processado.

transdução O movimento dos genes de uma bactéria doadora para uma receptora com um fago como vetor.

transdução especializada (restrita) A situação na qual um fago particular irá transduzir apenas regiões específicas do cromossomo bacteriano.

transdução generalizada A habilidade de alguns fagos em transduzir qualquer gene no cromossomo bacteriano.

transferência de Southern A transferência de fragmentos de DNA separados eletroforicamente de um gel para uma folha absorvente de papel; essa folha é então imersa em uma solução contendo uma sonda marcada que irá ligar-se ao fragmento de interesse.

transferência infecciosa A rápida transmissão de plasmídeos livres (mais quaisquer genes cromossômicos que eles possam carregar) das células doadoras para receptoras em uma população de bactérias.

transferência Northern A transferência de moléculas de RNA separadas eletroforicamente de um gel para uma folha absorvente, a qual é então imersa em uma sonda marcada que irá ligar-se ao RNA de interesse.

transferência Western Uma membrana levando uma impressão de proteína separada por eletroforese. Pode ser sondada com um anticorpo marcado para detectar uma proteína específica.

transformação A modificação direta de um genoma pela aplicação externa de DNA de uma célula com genótipo diferente.

transformação dupla Transformação simultânea por dois marcadores doadores diferentes.

transgene Um gene que foi modificado pela aplicação de técnicas de DNA recombinante e reintroduzido no genoma pela transformação da linhagem germinativa.

transição Um tipo de substituição de um par de nucleotídeos no qual uma purina substitui outra purina ou uma pirimidina substitui outra pirimidina, por exemplo, G-C para A-T.

transição alostérica Uma mudança de uma conformação de uma proteína para outra.

translocação A relocação de um segmento cromossômico em uma posição diferente no genoma.

translocação insercional A inserção de um segmento de um cromossomo em outro não-homólogo.

translocação recíproca Uma translocação na qual parte de um cromossomo é trocada com uma parte de um cromossomo não-homólogo separado.

transpor Mover-se de um local no genoma para outro; diz-se de um elemento genético móvel.

transposase Uma enzima codificada por elementos de transposição que sofrem uma transposição conservativa.

transposição Um processo pelo qual elementos genéticos movem-se de um local no genoma para outro.

transposição conservativa Um mecanismo de transposição que move um elemento móvel para um novo local no genoma quando o remove de sua localização anterior.

transposição replicativa Um mecanismo de transposição que gera um novo elemento de inserção integrado em outra parte no genoma deixando o elemento original em seu sítio original de inserção.

transposon (Tn) Um pedaço móvel de DNA que é flanqueado por sequências repetidas terminais e tipicamente possuem genes que codificam funções de transposição. Os transposons bacterianos podem ser simples ou compostos.

transposon composto Um tipo de elemento de transposição bacteriana contendo uma variedade de genes que residem entre dois elementos de sequência de inserção (IS) quase idênticos.

transposon de DNA *Veja elemento DNA.*

transposon simples Um tipo de elemento de transposição bacteriana contendo uma variedade de genes que residem entre sequências curtas de repetições invertidas.

transversão Um tipo de substituição de par de nucleotídeos no qual uma pirimidina substitui uma purina ou vice-versa, por exemplo, G-C por T-A.

triagem Um procedimento de mutagênese pelo qual essencialmente toda a prole mutada é recuperada e avaliada individualmente pelo fenótipo mutante; geralmente, o fenótipo desejado é marcado de algum modo para permitir sua detecção.

triagem genética *Veja triagem.*

trifosfato de adenosina *Veja ATP.*

trinca Três pares de nucleotídeos que constituem um códon.

triplóide Uma célula tendo três conjuntos cromossômicos ou um organismo composto de tais células.

trissômico Basicamente, um diplóide com um cromossomo extra de um tipo, produzindo um número de cromossomos da forma $2n + 1$.

trivalente Refere-se ao arranjo de pareamento meiótico de três homólogos em um triplóide ou trissômico.

tRNA (RNA transportador) Uma classe de pequenas moléculas de RNA que levam aminoácidos específicos para o ribossomo no curso da tradução; um aminoácido é inserido na cadeia polipeptídica crescente quando o anticódon do tRNA correspondente faz par com um códon no mRNA que está sendo traduzido.

tRNA carregado Uma molécula de RNA transportador com um aminoácido ligado à sua ponta 3'. Também chamado de aminoacil-tRNA.

tRNA isoacceptores Os vários tipos de moléculas de tRNA que levam um aminoácido específico.

U *Veja uracil; uridina.*

UAS *Veja seqüência de ativação antecedente.*

ubiquitina Uma proteína que, quando ligada como uma cadeia de várias cópias a outra proteína, marca essa proteína para degradação pela protease chamada 26S proteassomo. A adição de ubiquitinas isoladas a uma proteína pode mudar as interações proteína-proteína, como no caso de PCNA e polimerases de desvio (*bypass*).

ubiquitinização O processo de adição de ubiquitina a uma proteína marcada para degradação.

u.m. *Veja unidade de mapa.*

unidade de mapa (u.m.) A “distância” entre dois pares de genes ligados onde 1% dos produtos do meio é de recombinantes; uma unidade de distância em um mapa de ligação.

unidade de mapa genético (u.m.) A distância no mapa cromossômico correspondente a 1% de frequência de recombinantes.

univalente Um cromossomo meiótico não-pareado, como é em geral encontrado em trissômicos e triplóides.

universo A população extremamente grande da qual é colhida uma amostra para fins estatísticos.

uracil (U) Uma base pirimidínica no RNA no lugar da timina encontrada no DNA.

uridina (U) Um nucleosídeo tendo uracil como sua base.

UTR *Veja região não-traduzida 3'; região não-traduzida 5'.*

valor C O conteúdo de DNA de um genoma haplóide.

valor Lod Estatística cumulativa resumindo evidências que apóiam um valor de ligação específico.

valor médio parental A média dos valores de um fenótipo quantitativo para dois genitores específicos.

variação As diferenças entre os pais e sua prole ou entre membros individuais de uma população.

variação contínua Variação mostrando uma gama não-quebrada de valores fenotípicos.

variação de seqüência de DNA neutro Variação na seqüência de DNA que não está sob seleção natural.

variação descontínua Variação que tem classes distintas de fenótipos para determinada característica.

variação quantitativa A existência de uma gama de fenótipos de uma característica específica, diferindo em grau e não por diferenças qualitativas distintas.

variança Uma medida da variação ao redor da classe central de uma distribuição; o desvio quadrado médio das observações de seu valor médio.

variança ambiental A variança devida à variação ambiental.

variança de dominância Variança genética em um só locus atribuível à dominância de um alelo em relação a outro.

variança genética A variança fenotípica associada à diferença média em fenótipo entre genótipos diferentes.

variança genética aditiva Variança genética associada aos efeitos médios de substituir um alelo por outro.

variante Um organismo individual que é reconhecidamente diferente de um tipo padrão arbitrário nessa espécie.

variação A ocorrência dentro de um tecido de setores com fenótipos diferentes.

variação de efeito de posição (PEV) Variação causada pela inativação de um gene em algumas células por sua justaposição anormal com a heterocromatina.

vetor *Veja vetor de clonagem.*

vetor de clonagem Em clonagem, o plasmídeo ou cromossomo do fago usado para levar o segmento de DNA clonado.

vetor de expressão Um vetor com as apropriadas regiões reguladoras bacterianas situadas a 5' do sítio de inserção, permitindo a transcrição e tradução de uma proteína exógena na bactéria.

vetor de transporte Um vetor (p. ex., um plasmídeo) construído de tal modo que pode se replicar em pelo menos duas espécies de hospedeiros diferentes, permitindo que um segmento de DNA seja testado ou manipulado de vários modos na célula.

via de RNAi O mecanismo de interferência de RNA, pelo qual a presença de RNA bifilamentares leva à formação de pequenos RNA de interferência (siRNA) que guiam o complexo RISC para clivar mRNA complementares.

via desenvolvimental Uma cadeia de eventos moleculares que ocorrem em um conjunto de células equivalentes e produzem destinos diferentes entre essas células.

viabilidade A probabilidade de que um zigoto sobreviva e se desenvolva em um organismo adulto.

vigor híbrido Uma situação na qual uma F_1 é maior ou mais sadia que suas duas linhagens parentais puras.

vírus Uma partícula que consiste em núcleos e proteínas que devem infectar uma célula viva para se replicar e reproduzir.

VNTR *Veja número variável de repetições em tandem.*

WGS *Veja seqüenciamento shotgun de genoma inteiro.*

xeroderma pigmentoso (XP) Um distúrbio causado por mutações no sistema de reparo por excisão de nucleotídeo acoplado a transcrição que leva ao freqüente desenvolvimento de cânceres de pele.

XP *Veja xeroderma pigmentoso.*

YAC *Veja cromossomo artificial de levedura.*

zigoto A célula formada pela fusão de um ovócito e um espermatozóide; a única célula diplóide que irá dividir-se mitoticamente para criar um organismo diplóide diferenciado.

Respostas de Problemas Seleccionados

Capítulo 1

2. O DNA determina todos os atributos específicos de uma espécie (forma, tamanho, características comportamentais, processos bioquímicos etc.) e estabelece os limites da possível variação que é induzida ambientalmente.

5. Se o DNA é uma dupla hélice, $A = T$ e $G = C$ e $A + T + C + G = 100\%$. Se $T = 15\%$, então $C = [100 - 15(2)]/2 = 35\%$.

6. Se o DNA é uma dupla hélice, $G = C = 24\%$ e $A = T = 26\%$.

10. a. Sim. Porque $A = T$ e $G = C$, a equação $A + C = G + T$ pode ser reescrita como $T + C = C + T$ substituindo termos iguais.

b. Sim. A percentagem de purinas será igual à percentagem de pirimidinas no DNA bifilamentar.

15. A eletroforese separa moléculas de DNA por tamanho. Quando o DNA é cuidadosamente isolado de *Neurospora* (que tem sete cromossomos diferentes), devem ser produzidas sete bandas com o uso dessa técnica. Similarmente, a ervilha tem sete cromossomos diferentes e produzirá sete bandas (cromossomos homólogos migram juntos como uma só banda). A mosca doméstica tem seis cromossomos diferentes e deve produzir seis bandas.

18. a. Uma planta incapaz de produzir pigmento vermelho teria pétalas azuis.

b. Uma planta incapaz de produzir o pigmento azul teria pétalas vermelhas.

c. Uma planta incapaz de produzir ambos os pigmentos, azul e vermelho, teria pétalas brancas.

Capítulo 2

2. PFGE separa moléculas de DNA por tamanho. Quando o DNA é cuidadosamente isolado de *Neurospora* (que tem sete cromossomos diferentes), devem ser produzidas sete bandas com o uso dessa técnica. Similarmente, a ervilha tem sete cromossomos diferentes e produzirá sete bandas (cromossomos homólogos irão comigrar como uma única banda).

5. A função principal da mitose é gerar duas células-filhas geneticamente idênticas à célula parental original.

9. À medida que as células se dividem mitoticamente, cada cromossomo consiste em cromátides-irmãs idênticas que são separadas para formar células-filhas geneticamente idênticas. Embora a segunda divisão da meiose pareça um processo similar, as cromátides-irmãs provavelmente são diferentes umas das outras. A recombinação nos primeiros estágios da meiose trocou regiões de DNA entre cromossomos-irmãos e não-irmãos, de modo que as células-filhas dessa divisão tipicamente não são geneticamente idênticas.

13. Sim. Metade de nossa constituição genética é derivada de cada um dos genitores etc.

17. (5) Sinapse (pareamento de cromossomos)

22. A proporção da prole é de aproximadamente 3:1, indicando a clássica reprodução entre heterozigotos. Porque Black (B) é dominante em relação a white (b),

Genitores: $B/b \times B/b$

Prole: 3 pretos : 1 branco (1 B/B : 2 B/b : 1 b/b)

26. O fato de que cerca de metade da prole é de mutantes sugere que a mutação que resulta em três cotilédones é dominante, e o mutante original era heterozigoto. Se C = alelo mutante e c = alelo tipo selvagem, o cruzamento é o seguinte:

P $C/c \times c/c$

F₁ C/c três cotilédones

c/c dois cotilédones

31. p (criança tem galactosemia) = p (John é G/g) $\times p$ (Martha é G/g) $\times p$ (ambos os genitores passaram g para a criança) = $(2/3)(1/4)(1/4) = 2/48 = 1/24$

37. a. O distúrbio parece ser dominante porque todos os indivíduos afetados têm um genitor afetado. Se a característica fosse recessiva, então I-1, II-2, III-1 e III-8 teriam todos que ser portadores (heterozigotos para o alelo raro).

b. Com a suposição de dominância, os genótipos são

I: $d/d, D/d$

II: $D/d, d/d, D/d, d/d$

III: $d/d, D/d, d/d, D/d, d/d, d/d, D/d, d/d$

IV: $D/d, d/d, D/d, d/d, d/d, d/d, D/d, d/d$

c. A probabilidade de um filho afetado (D/d) é igual a $1/2$ e a probabilidade de um filho não-afetado (d/d) é igual a $1/2$. Portanto, a chance de ter quatro filhos não-afetados (como cada filho é um evento independente) é: $(1/2) \times (1/2) \times (1/2) \times (1/2) = 1/16$.

43. a. Os filhos herdam o cromossomo X de suas mães. A mãe tem lobos na orelha; o filho não. Se o achado de lobos na orelha é dominante e o alelo para falta de lobos é recessivo, então a mãe pode ser heterozigota para essa característica e o gene pode ser ligado ao X.

b. Não é possível, com os dados fornecidos, decidir qual alelo é dominante. Se a falta de lobos na orelha é dominante, então o pai seria heterozigoto e o filho teria 50% de herdar o alelo dominante "falta de lobos na orelha". Se a falta de lobos na orelha é recessivo, então a característica pode ser autossômica ou ligada ao X, mas, em qualquer dos casos, a mãe seria heterozigota.

47. Seja H = a hipofosfatemia e h = normal. A reprodução é $H/Y \times h/h$, produzindo H/h (mulheres) e h/Y (homens). A resposta é 0%.

52. a. $X^C/X^c, X^c/X^c$

b. p (daltônica) $\times p$ (homem) = $(1/2)(1/2) = 1/4$

c. As meninas serão 1 normal (X^C/X^c): 1 daltônica (X^c/X^c)

d. A reprodução é $X^C/X^c \times X^c/Y$, produzindo 1 normal : 1 daltônico para ambos os sexos.

60. a. O padrão de herança para cabelo ruivo sugerido por esse heredograma é recessivo porque a maioria das pessoas ruivas têm genitores sem essa característica.

b. A observação dos que estão a nossa volta faz o alelo parecer um tanto raro.

64. Note que apenas os homens são afetados, e que, em todos menos um caso, a característica pode ser rastreada pelo lado feminino. Entretanto, há um exemplo de um homem afetado tendo filhos afetados. Se a característica é ligada ao X, a esposa desse homem deve ser uma portadora, sugerindo que o distúrbio seja causado por um alelo autossômico dominante com expressão limitada aos homens, dependendo do quanto essa característica é rara na população geral.

Capítulo 3

2. O genótipo das células-filhas será idêntico ao da célula original: (f) $A/a; B/b$.

7. A mitose produz células-filhas tendo o mesmo genótipo que a célula original: $A/a; B/b; C/c$.

10. Seus filhos terão que herdar o satélite contendo 4 (probabilidade = $1/2$), a coloração anormal 7 (probabilidade = $1/2$) e o cromossomo Y (probabilidade = $1/2$). Para herdar todos três, a probabilidade é $(1/2)(1/2)(1/2) = 1/8$.

15. Com a suposição de distribuição independente e simples relações dominante-recessivo de todos os genes, o número de classes genótípicas esperadas da autofecundação de uma planta heterozigota para n pares de genes é 3^n , e o número de classes fenotípicas esperadas é 2^n .

18. a. e b. O cruzamento 2 indica que púrpura (G) é dominante em relação a verde (g), e o cruzamento 1 indica que cut (P) é dominante em relação a potato (p).

Cruzamento 1: $G/g; P/p \times g/g; P/p$ Existem 3 cut : 1 potato e 1 púrpura : 1 verde.

Cruzamento 2: $G/g; P/p \times G/g; p/p$ Três são púrpura, 1 verde e 1 cut : 1 potato.

Cruzamento 3: $G/G; P/p \times g/g; P/p$ Não existe verde, e existem 3 cut : 1 potato.

Cruzamento 4: $G/g; P/P \times g/g; p/p$ Não há potato, e existe 1 púrpura : 1 verde.

Cruzamento 5: $G/g; p/p \times g/g; P/p$ Há 1 cut : 1 potato, e existe 1 púrpura : 1 verde.

23. Os cruzamentos são

Cruzamento 1: fêmea pára-começa \times macho tipo selvagem \rightarrow toda a prole pára-começa

Cruzamento 2: fêmea tipo selvagem \times macho pára-começa \rightarrow toda a prole tipo selvagem

O mtDNA é herdado apenas da “fêmea” em *Neurospora*.

29. a. Devem existir nove classes correspondendo a 0, 1, 2, 3, 4, 5, 6, 7, 8 “doses”.

b. Devem existir treze classes correspondendo a 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 “doses”.

34. As plantas da prole herdam apenas cpDNA normal (coluna 1); apenas cpDNA mutante (coluna 2), ou ambas (coluna 3). Para obter cpDNA, vistos nas colunas 1 e 2, os cloroplastos tiveram que ser segregados.

39. a e b. Comece com quaisquer duas das três linhagens e as cruze. Se, por exemplo, você começar com a/a ; B/B ; $C/C \times A/A$; b/b ; C/C , toda a prole seria A/a ; B/b ; C/C . Cruzando duas delas, teríamos

9 $A/-$; $B/-$; C/C

3 a/a ; $B/-$; C/C

3 $A/-$; b/b ; C/C

1 a/a ; b/b ; C/C

O genótipo a/a ; b/b ; C/C tem dois dos genes em um estado homozigoto recessivo, e é encontrado em 1/16 da prole. Se esse genótipo fosse cruzado com A/A ; B/B ; c/c , toda a prole seria A/a ; B/b ; C/c . O cruzamento de duas delas (ou “autofecundação”) levaria a uma proporção 27 : 9 : 9 : 9 : 3 : 3 : 3 : 1, e 1/64 da prole seria a desejada a/a ; b/b ; c/c .

Existem várias vias diferentes para obter a/a ; b/b ; c/c , mas a destacada requer apenas quatro cruzamentos.

44. a. Em uma célula diplóide, esperam-se dois cromossomos (um par de homólogos), cada um tendo um único locus de radioatividade.

b. Esperam-se muitas regiões de radioatividade dispersas pelos cromossomos. O número exato e o padrão dependeriam da sequência específica em questão, e onde e com que frequência está presente no genoma.

c. As múltiplas cópias dos genes para RNA ribossômico são organizadas em grandes arranjos em tandem chamados de organizadores nucleolares. Portanto, esperam-se áreas mais largas de radioatividade comparadas com a da parte a. O número dessas regiões seria igual ao número de organizadores nucleolares presentes no organismo.

d. Espera-se que cada ponta de cromossomo seja marcada por DNA telomérico.

e. As múltiplas repetições desse DNA heterocromático estão organizadas em grandes arranjos em tandem. Portanto, espere áreas mais largas de radioatividade comparadas com a da parte a. Também pode haver mais de uma área no genoma da mesma repetição simples.

48. a. Seja B = braquidactilia, b = normal, T = sensível e t = insensível. Os genótipos do casal são B/b ; T/t para o macho e b/b ; T/t para a fêmea.

b. Para todas as quatro crianças terem braquidactilia, $p = (1/2)^4 = 1/16$.

c. Para nenhuma das quatro crianças ter braquidactilia, $p = (1/2)^4 = 1/16$.

d. Para todas serem sensíveis, $p = (3/4)^4 = 81/256$.

e. Para todas serem insensíveis, $p = (1/4)^4 = 1/256$.

f. Para que todas tenham braquidactilia e sejam sensíveis, $p = (1/2 \times 3/4)^4 = 81/4.096$.

g. A probabilidade de ser sensível e não ter braquidactilia é de $1 -$ (a probabilidade de ser sensível e ter braquidactilia), ou $1 - (1/2 \times 3/4) = 5/8$. A probabilidade de que todas as crianças não sejam braquidáctilos sensíveis é de $(5/8)^4 = 625/4.096$.

h. A probabilidade de que pelo menos uma tenha braquidactilia e seja sensível é de $1 -$ (a probabilidade de nenhuma ter braquidactilia e ser sensível), ou $1 - (5/8)^4$.

Capítulo 4

2. P $A d/A d \times a D/a D$
 F_1 $A d/a D$
 F_2 1 $A d/A d$ fenótipo: $A d$
 2 $A d/a D$ fenótipo: $A D$
 1 $a D/a D$ fenótipo: $a D$

5. Como só os tipos parentais são recuperados, os dois genes devem estar fortemente ligados, e a recombinação deve ser muito rara. Sabendo quanto de prole estamos procurando, teríamos uma indicação do quanto os genes estão próximos.

10. a. Os três genes estão ligados.

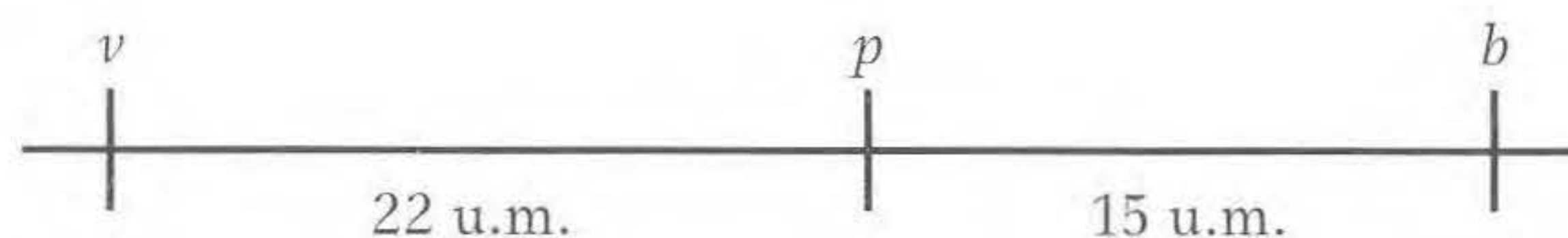
b. Uma comparação dos parentais (mais frequentes) com os crossings duplos (menos frequentes) revela que a ordem dos genes é $v p b$. Havia

2.200 recombinantes entre v e p , e 1.500 entre p e b . A fórmula geral para unidades de mapa é

$$\text{u.m.} = 100\% (\text{número de recombinantes}) / \text{número total da prole}$$

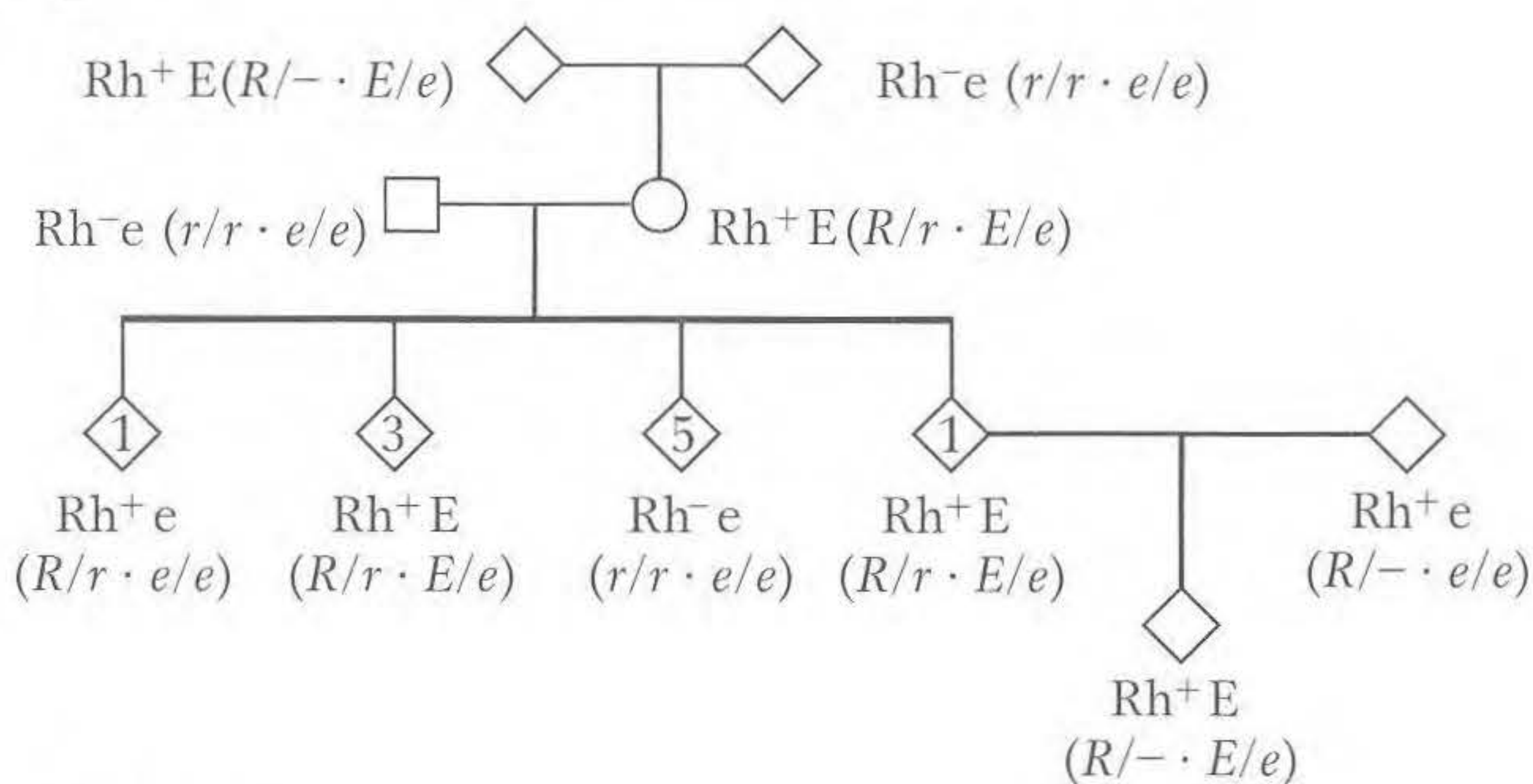
Portanto, as unidades de mapa entre v e $p = 100\%(2.200)/10.000 = 22 \text{ u.m.}$, e as unidades de mapa entre p e $b = 100\%(1.500)/10.000 = 15 \text{ u.m.}$

O mapa é



$$\begin{aligned} \text{c. } I &= 1 - \text{crossings duplos observados} / \text{crossings duplos esperados} \\ &= 1 - 132 / (0,22)(0,15)(10.000) \\ &= 1 - 0,4 = 0,6 \end{aligned}$$

16. a.



b. Sim.

c. Dominante.

d. Como desenhado, o heredograma sugere ligação. Se não ligado, espera-se que os fenótipos das 10 crianças estejam em uma proporção 1 : 1 : 1 : 1 de $Rh^+ E$, $Rh^+ e$, $Rh^- E$ e $Rh^- e$. Existem de fato cinco $Rh^- e$, quatro $Rh^+ E$ e um $Rh^+ e$. Se ligados, este último fenótipo representaria um recombinante, e a distância entre os dois genes seria de $100\% (1/10) = 10 \text{ u.m.}$ Entretanto, não existem dados suficientes para apoiar fortemente essa conclusão.

22. a. Se os genes não estão ligados, o cruzamento é

P hyg/hyg ; $her/her \times hyg^+/hyg^+$; her^+/her^+

F_1 hyg^+/hyg ; $her^+/her \times hyg^+/hyg$; her^+/her

F_2 9/16 $hyg^+/-$; $her^+/-$

3/16 $hyg^+/-$; her/her

3/16 hyg/hyg ; $her^+/-$

1/16 hyg/hyg ; her/her

Assim, apenas 1/16 (ou 6,25%) das sementes devem germinar.

b. e c. Não. Germinaram mais que o dobro das sementes esperadas; logo, suponha que os genes estão ligados. O cruzamento então é

P $hyg her/hyg her \times hyg^+ her^+/hyg^+ her^+$

F_1 $hyg^+ her^+/hyg her \times hyg^+ her^+/hyg her$

F_2 13% $hyg her/hyg her$

Como essa classe representa a combinação de dois cromossomos parentais, é igual a

$$p(hyg her) \times p(hyg her) = (1/2 \text{ parentais})^2 = 0,13$$

e

$$\text{parentais} = 0,72$$

logo

$$\text{recombinantes} = 1 - 0,72 = 0,28$$

Portanto, um cruzamento-teste de $hyg^+ hyg^+/hyg her$ dará

36% $hyg^+ her^+/hyg her$

36% $hyg her/hyg her$

14% $hyg^+ her/hyg her$

14% $hyg her^+/hyg her$

e 36% da prole crescerá (a classe $hyg her/hyg her$).

27. A fórmula para este problema é $f(i) = e^{-m} m^i / i!$, onde $m = 2$ e $i = 0, 1$, ou 2

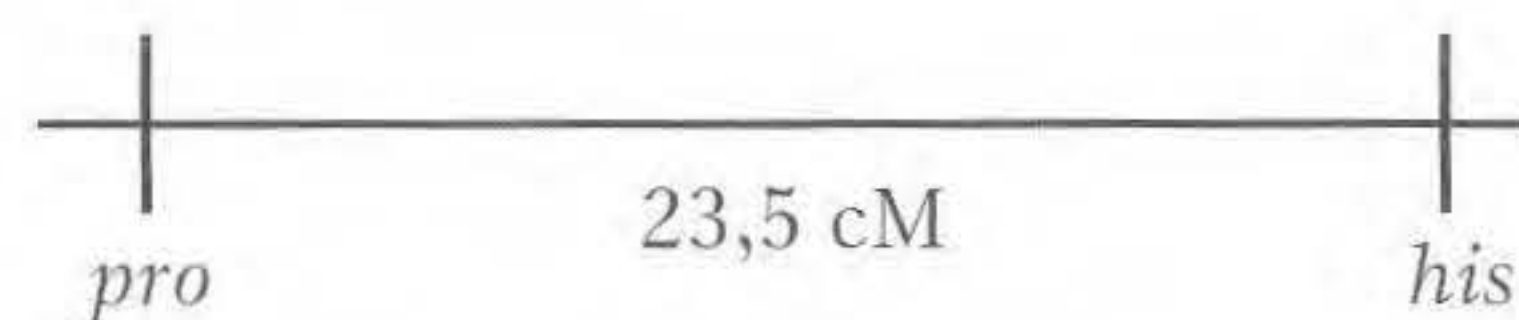
a. $f(0) = e^{-2} 2^0 / 0! = e^{-2} = 0,135$, ou 13,5%.

b. $f(1) = e^{-2} 2^1 / 1! = e^{-2} (2) = 0,27$, ou 27%.

c. $f(2) = e^{-2} 2^2 / 2! = e^{-2} (2) = 0,27$, ou 27%.

33. a. O cruzamento foi $pro \times his$, o que dá a primeira classe de tetrade DNP (6 ditipos não-parentais), a segunda classe de tetrade T (82 tetratipos) e a terceira classe de tetrade DP (112 ditipos parentais). Quando $DP \gg DNP$, você sabe que os dois genes estão ligados.

b. A distância de mapa pode ser calculada usando a fórmula $FR = [DNP + (1/2)T]100\%$. Nesse caso, a frequência é DNP 6/200, ou 3%, e a frequência de T é 82/200, ou 41%. A distância de mapa entre esses dois loci é, portanto, 23,5 cM.



c. Para corrigir para vários crossings, pode ser usada a fórmula de Perkins. Assim, a distância de mapa = $(T + 6DNP)50\%$, ou $(0,41 + 0,18)50\% = 29,5$ cM.

37. a. O cruzamento é $W e F / W e F \times w e f / w e f$ e a F_1 é $W e F / w e f$. A prole que era $ww ee ff$ de um cruzamento-teste dessa F_1 deve ter herdado um dos cromossomos recombinantes duplo crossing ($w e f$). Com a suposição de não-interferência, a percentagem esperada de crossings duplos é $8\% \times 24\% = 1,92\%$, metade da qual é 0,96%.

b. Para obter uma prole $ww ee ff$ de uma autofecundação dessa F_1 , é necessária a herança independente de dois cromossomos recombinantes duplos $w e f$. Suas chances de ocorrer, com base na resposta da parte a deste problema, são $0,96 \times 0,96 = 0,009\%$.

43. O cruzamento é $A/A \cdot B/B \times a/a \cdot b/b$ e, então, o diíbrido resultante de F_1 é submetido a cruzamento-teste. Portanto, para fins de cálculo, a prole do cruzamento-teste que é $A/a \cdot B/b$ ou $a/a \cdot b/b$ deve ser considerada "parental", e a prole que é $A/a \cdot b/b$ ou $a/a \cdot B/b$ deve ser considerada de "recombinantes".

O quadro seguinte mostra as proporções esperadas de genótipos parentais (P) e recombinantes (R) para distribuição independente e os quatro valores de FR deste problema.

	FR			
	0,5	0,3	0,2	0,1
P	0,25	0,35	0,4	0,45
P	0,25	0,35	0,4	0,45
R	0,25	0,15	0,1	0,05
R	0,25	0,15	0,1	0,05

Na ordem de nascimento, a prole é

$A/a \cdot B/b$	parental
$a/a \cdot b/b$	parental
$A/a \cdot B/b$	parental
$A/a \cdot b/b$	recombinante
$a/a \cdot b/b$	parental
$A/a \cdot B/b$	parental
$a/a \cdot B/b$	recombinante

A probabilidade de obter esses resultados se os dois genes não estiverem ligados (FR de 50%) será $0,25 \times 0,25 \times 0,25 \times 0,25 \times 0,25 \times 0,25 \times B = 0,000061B$ (onde B = número de possíveis ordens de nascimento para cinco indivíduos parentais e dois recombinantes).

a. A probabilidade de obter esses resultados se os dois genes estão ligados (FR de 10%) será $0,45 \times 0,45 \times 0,45 \times 0,05 \times 0,45 \times 0,45 \times B = 0,000046B$. A proporção dessa probabilidade em relação ao resultado esperado se os genes não estiverem ligados é $0,000046B/0,000061B = 0,756$. O logaritmo dessa proporção é o valor Lod, ou, nesse caso, $Lod = \log(0,756) = -0,12$.

b. A probabilidade de obter esses resultados se os dois genes estão ligados (FR de 20%) será $0,40 \times 0,40 \times 0,40 \times 0,10 \times 0,40 \times 0,40 \times B = 0,000102B$. A proporção dessa probabilidade em relação ao resultado esperado se os genes não estiverem ligados é 1,68, e o valor Lod será 0,22.

c. A probabilidade de obter esses resultados se os dois genes estão ligados (FR de 30%) será 0,000118B; a proporção será 1,68, e o valor Lod será 0,29.

48. O dados para cada um dos cruzamentos-teste de três pontos podem ser usados para determinar a ordem dos genes percebendo que as classes recombinantes mais raras são o resultado de eventos de crossing duplos. Uma comparação desses cromossomos com os tipos "parentais" revela que os alelos que mudaram representam o gene no meio.

Por exemplo, no conjunto de dados 1, os fenótipos mais comuns ($+++$ e $a b c$) representam as combinações alélicas parentais. Uma comparação desses fenótipos com os fenótipos mais raros desse conjunto de dados ($+ b c$ e $a ++$) indica que o gene a é recombinante e deve estar no meio. A ordem dos genes é $b a c$.

Para o conjunto de dados 2, $+ b c$ e $a ++$ (parentais) devem ser comparados com $+++$ e $a b c$ (os recombinantes mais raros) para indicar que o gene a está no meio. A ordem dos genes é $b a c$.

Para o conjunto de dados 3, compare $+ b + e a + c$ com $a b + e ++ c$, que dá a ordem de genes $b a c$.

Para o conjunto de dados 4, compare $++ c$ e $a b +$ com $+++$ e $a b c$, que dá a ordem de genes $a c b$.

Para o conjunto de dados 5, compare $+++$ e $a b c$ com $++ c$ e $a b +$, que dá a ordem de genes $a c b$.

54. a. O cruzamento 1 reduz-se a

P	$A/A \cdot B/B \cdot D/D \times a/a \cdot b/b \cdot d/d$
F_1	$A/a \cdot B/b \cdot D/d \times a/a \cdot b/b \cdot d/d$

A prole do cruzamento-teste indica que esses três genes estão ligados (OC = crossing, DCO = duplo crossing).

Prole de	$A B D$	316	parental
cruzamento-teste	$a b d$	314	parental
	$A B d$	31	CO B-D
	$a b D$	39	CO B-D
	$A b d$	130	CO A-B
	$a B D$	140	CO A-B
	$A b D$	17	DCO
	$a B d$	13	DCO

$$A-B: 100\%(130 + 140 + 17 + 13)/1.000 = 30 \text{ u.m.}$$

$$B-D: 100\%(31 + 39 + 17 + 13)/1.000 = 10 \text{ u.m.}$$

O cruzamento 2 reduz-se a

P	$A/A \cdot C/C \cdot E/E \times a/a \cdot c/c \cdot e/e$
F_1	$A/a \cdot C/c \cdot E/e \times a/a \cdot c/c \cdot e/e$

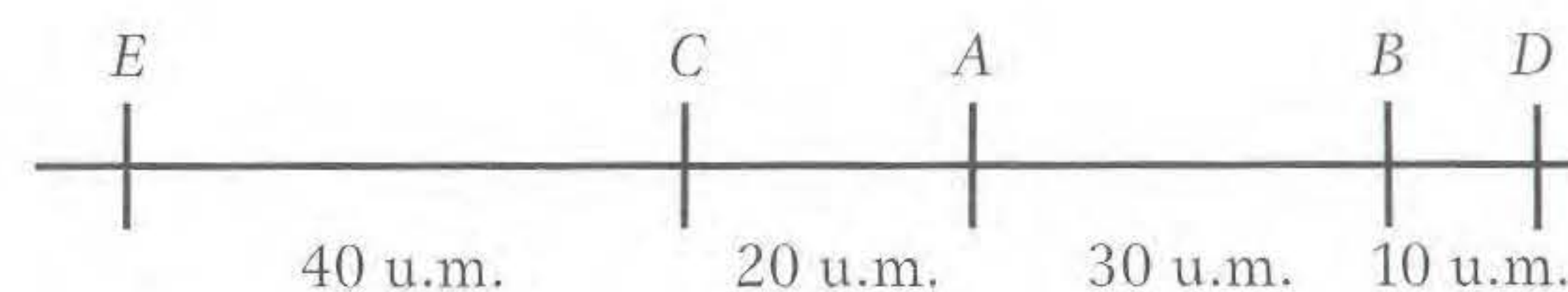
A prole do cruzamento-teste indica que esses três genes estão ligados.

Prole de	$A C E$	243	parental
cruzamento-teste	$a c e$	237	parental
	$A c e$	62	CO A-C
	$a C E$	58	CO A-C
	$A C e$	155	CO C-E
	$a c E$	165	CO C-E
	$a C e$	46	DCO
	$A c E$	34	DCO

$$A-B: 100\%(62 + 58 + 46 + 34)/1.000 = 20 \text{ u.m.}$$

$$B-D: 100\%(155 + 165 + 46 + 34)/1.000 = 40 \text{ u.m.}$$

O mapa que acomoda todos os dados é



b. Interferência (I) = $1 - [(DCO \text{ observado})/(DCO \text{ esperado})]$

Para o cruzamento 1: $I = 1 - \{30/[(0,30)(0,10)(1.000)]\} = 1 - 1 = 0$, sem interferência

Para o cruzamento 2: $I = 1 - \{80/[(0,20)(0,40)(1.000)]\} = 1 - 1 = 0$, sem interferência

59. a e b. Os dados apóiam a distribuição independente de dois genes (sejam *arg1* e *arg2*). O cruzamento torna-se $arg1 ; arg2^+ \times arg1^+ ; arg2$ e as tetrades resultantes são

4 : 0 (DP)	3 : 1 (T)	2 : 2 (DNP)
$arg1 ; arg2^+$	$arg1 ; arg2^+$	$arg1 ; arg2$
$arg1 ; arg2^+$	$arg1^+ ; arg2$	$arg1 ; arg2$
$arg1^+ ; arg2$	$arg1 ; arg2$	$arg1^+ ; arg2^+$
$arg1^+ ; arg2$	$arg1^+ ; arg2^+$	$arg1^+ ; arg2^+$

Como $DP = DNP$, os genes não estão ligados.

Capítulo 5

1. Uma linhagem Hfr tem o fator de fertilidade F integrado ao cromossomo. Uma linhagem F^+ tem o fator de fertilidade livre no citoplasma. Uma linhagem F^- não tem o fator de fertilidade.

5. Embora os experimentos de reprodução interrompida produzam a ordem dos genes, isso será relativo apenas a marcadores distantes. Assim, a mutação não pode ser situada com precisão com essa técnica. A transdução generalizada produzirá informações com relação a marcadores bem próximos, o que torna uma escolha pobre para os experimentos iniciais devido à intensa quantidade de triagens que teriam que ser feitas. Juntas, as duas técnicas permitem, primeiro, localizar o mutante (reprodução interrompida) e, segundo, a determinação precisa da localização do mutante (transdução generalizada) dentro da região geral.

10. A melhor explicação é que o fator F integrado da alça de Hfr do cromossomo bacteriano é anormal e agora é um F' que contém o gene pro^+ . Esse F' é rapidamente transferido para células F^- , convertendo-as em pro^+ (e F^+).

15. O número esperado de recombinantes duplos é (0,01) (0,002) (100.000) = 2. Interferência = $1 - (\text{crossings duplos observados} / \text{crossings duplos esperados}) = 1 - 5/2 = -1,5$. Por definição, a interferência é negativa.

19. a. Esse processo parece ser transdução especializada. Ela é caracterizada pela transdução de marcadores especializados baseados na posição da integração do profago. Apenas os genes perto do sítio de integração são possíveis candidatos à má incorporação em partículas de fago que, então, levam esse DNA para bactérias receptoras.

b. Os únicos meios que suportam o crescimento de colônias eram os que não tinham cisteína ou leucina. Esses meios são selecionados para transdutantes cis^+ ou leu^+ e indicam que o profago está situado na região *cis-leu*.

24. Não. Loci proximamente ligados seriam esperados sendo co-transduzidos; quanto maior a frequência de co-transdução, mais próximos estão os genes. Como apenas 1 das 858 $metE^+$ também era $pirD^+$, os genes não estão proximamente ligados. A única $metE^+$ $pirD^+$ pode ser o resultado de co-transdução, ou uma mutação espontânea de $pirD$ para $pirD^+$, ou o resultado de co-infecção por dois fagos transduzidos separadamente.

29. a. Para determinar quais genes estão próximos, compare as frequências dos duplos transformantes. O teste par a par dá valores baixos sempre que B é incluído, mas taxas relativamente altas quando qualquer droga menos B é incluída. Esse achado sugere que o gene para resistência a B não está próximo dos outros três genes.

b. Para determinar a ordem relativa dos genes para resistência a A, C e D, compare as frequências do duplo e do triplo transformantes. A frequência da resistência a AC é aproximadamente a mesma que a resistência a ACD, que sugere fortemente que D está no meio. Além disso, a frequência de co-resistência a AD é maior que AC (sugerindo que o gene para resistência a A está mais perto de D do que de C) e a frequência de CD é maior que AC (sugerindo que C está mais próximo de D do que de A).

33. Para isolar as partículas transdutoras especializadas do fago $\phi 80$ que leva lac^+ , os pesquisadores tiveram que lisogenizar a linhagem com $\phi 80$, induzir o fago com UV e, então, usar os lisados para transduzir uma linhagem Lac para Lac^+ . As colônias Lac^+ seriam então usadas para fazer um novo lisado, que teria que ser altamente enriquecido para o fago transdutor lac^+ .

Capítulo 6

2. Com a suposição de homozigose para o gene normal, a reprodução é $A/A \cdot b/b \times a/a \cdot B/B$. As crianças seriam normais, $A/a \cdot B/b$.

5. a. Vermelho	b. Púrpura
c. 9	$M_1/- ; M_2/-$ púrpura
3	$m_1/m_1 ; M_2/-$ azul
3	$M_1/- ; m_2/m_2$ vermelho
1	$m_1/m_1 ; m_2/m_2$ branco

d. Os alelos mutantes não produzem uma enzima funcional. Entretanto, suficiente enzima funcional deve ser produzida pelo único alelo tipo selvagem de cada gene para sintetizar níveis normais do pigmento.

9. a. O cruzamento original foi diíbrido. Tanto oval quanto púrpura devem representar um fenótipo incompletamente dominante.

b. Um cruzamento longo, púrpura \times oval, púrpura é o seguinte:

P $L/L ; R/R' \times L/L' ; R/R''$

F_1	$\frac{1}{2} L/L \times$	$\frac{1}{4} R/R$	$\frac{1}{8}$ longo, vermelho
		$\frac{1}{2} R/R'$	$\frac{1}{4}$ longo, púrpura
		$\frac{1}{4} R'/R'$	$\frac{1}{8}$ longo, branco
	$\frac{1}{2} L/L' \times$	$\frac{1}{4} R/R$	$\frac{1}{8}$ oval, vermelho
		$\frac{1}{2} R/R'$	$\frac{1}{4}$ oval, púrpura
		$\frac{1}{4} R'/R'$	$\frac{1}{8}$ oval, branco

12.

	Genitores	Filhos
a.	$AB \times O$	B
b.	$A \times O$	A
c.	$A \times AB$	AB
d.	$O \times O$	O

16. a. A proporção de sexos esperada é de 1:1.

b. A genitora era heterozigota para um alelo letal recessivo ligado ao X, que resultaria em 50% menos machos do que fêmeas.

c. Metade da prole feminina deve ser heterozigota para o alelo letal, e metade deve ser homozigota para o alelo não-letal. Individualmente, cruze as fêmeas F_1 e determine a proporção de sexo de sua prole.

19. a. As mutações são em dois genes diferentes porque o heterocáion é prototrófico (as duas mutações se complementam.)

b. $leu1^+ ; leu2^-$ e $leu1^- ; leu2^+$

c. Com distribuição independente, esperamos

$\frac{1}{4}$	$leu1^+ ; leu2^-$
$\frac{1}{4}$	$leu1^- ; leu2^+$
$\frac{1}{4}$	$leu1^- ; leu2^-$
$\frac{1}{4}$	$leu1^+ ; leu2^+$

23. a. P A/a (frizzle) $\times A/a$ (frizzle)

F_1 1 A/A (normal) : 2 A/a (frizzle) : 1 a/a (lanoso)

b. Se A/A (normal) é cruzado com a/a (lanoso), toda a prole seria A/a (frizzle).

25. A produção de prole preta de genitores albinos recessivos puros é possível se o albinismo resultar de mutações em dois genes diferentes. Se o cruzamento for

$A/A ; b/b \times a/a ; B/B$

toda a prole seria

$A/a ; B/b$

e eles teriam o fenótipo preto devido à complementação.

29. O genitor púrpura pode ser $A/a ; b/b$ ou $a/a ; B/b$ para essa resposta. Suponha que o genitor púrpura seja $A/a ; b/b$. O genitor azul deve ser $A/a ; B/b$.

32. O cruzamento é cinza \times amarelo, ou $A/- ; R/- \times A/- ; r/r$. A prole F_1 é

$\frac{3}{8}$ amarela	$\frac{1}{8}$ preta	$\frac{3}{8}$ cinza	$\frac{1}{8}$ branca
-----------------------	---------------------	---------------------	----------------------

Para a prole branca, ambos os genitores devem ter um alelo r e um alelo a . Agora o cruzamento pode ser reescrito como $A/a ; R/r \times A/a ; r/r$.

35. O cachorro original marrom é $w/w ; b/b$, e o cachorro branco original é $W/W ; B/B$. A prole F_1 é $W/w ; B/b$ e a prole F_2 é:

9 $W/- ; B/-$	branco
3 $w/w ; B/-$	preto
3 $W/- ; b/b$	branco
1 $w/w ; b/b$	marrom

39. Os heredogramas como este são muito comuns. Eles indicam falta de penetrância devido a epistasia ou efeitos ambientais. O indivíduo A deve ter o gene autossômico dominante.

41. a. Seja W^O = oval, W^S = afoiçado e W^R = redondo. Os três cruzamentos são

Cruzamento 1: $W^S/W^S \times W^R/Y \rightarrow W^S/W^R$ e W^S/Y

Cruzamento 2: $W^S/W^S \times W^S/Y \rightarrow W^S/W^R$ e W^R/Y

Cruzamento 3: $W^S/W^S \times W^O/Y \rightarrow W^O/W^S$ e W^S/Y

b. $W^O/W^S \times W^R/Y$

$\frac{1}{4} W^O/W^R$	fêmea oval
$\frac{1}{4} W^S/W^R$	fêmea afoiçada
$\frac{1}{4} W^O/Y$	macho oval
$\frac{1}{4} W^S/Y$	macho afoiçado

44. a. Os genótipos são:

P	$B/B ; i/i \times b/b ; I/I$	
F_1	$B/B ; I/i$	sem cerdas
F_2	$9 B/- ; I/-$	sem cerdas
	$3 B/- ; i/i$	retas
	$3 b/b ; I/-$	sem cerdas
	$1 b/b ; i/i$	curva

b. Os genótipos são: $B/b ; I/i \times B/b ; i/i$.

47. Existe um total de prole de 159 que deve ser distribuído em uma proporção de 9:3:3:1 se os dois genes se distribuem independentemente. Você pode ver que

Observado	Esperado
88 $P/- ; Q/-$	90
32 $P/- ; q/q$	30
25 $p/p ; Q/-$	30
14 $p/p ; q/q$	10

50. a. Está ocorrendo alimentação cruzada, onde o produto feito por uma linhagem difunde-se para outra linhagem e permite o crescimento da segunda linhagem.

b. Para que ocorra a alimentação cruzada, a linhagem que cresce deve ter um bloqueio que está presente mais antecedente na via metabólica que o bloqueio na linhagem da qual a linhagem crescente está obtendo o produto para crescimento.

c. Os dados sugerem que a via metabólica é



d. Sem algum triptofano, não haveria nenhum crescimento, e as células não teriam vivido o suficiente para gerar o produto que poderia se difundir.

52. a. A melhor explicação é que a síndrome de Marfan é herdada como uma característica autossômica dominante.

b. O heredograma mostra pleiotropia (múltiplas características afetadas) e expressividade variável (grau variável de expressão fenotípica) e expressividade variável (grau variável de fenótipo expresso).

c. A pleiotropia indica que o produto gênico é necessário em vários tecidos diferentes, órgãos ou processos. Quando o gene é mutante, todos os tecidos que precisam do produto gênico serão afetados. A expressividade variável de um fenótipo para determinado genótipo indica modificação de um ou mais outros genes, ruído aleatório ou efeitos ambientais.

55. a. Esse tipo de interação gênica é chamado de *epistasia*. O fenótipo de *e/e* é epistático aos fenótipos de $B/-$ ou b/b .

b. Os genótipos deduzidos são os seguintes:

I	1 ($B/b E/e$)	2 ($B/b E/e$)		
II	1 ($b/b E/e$)	2 ($B/b E/e$)	3 ($-/- e/e$)	4 ($b/b E/-$)
	5 ($B/b E/e$)	6 ($b/b E/e$)		
III	1 ($B/b E/-$)	2 ($-/b e/e$)	3 ($b/b E/-$)	4 ($B/b E/-$)
	5 ($b/b E/-$)	6 ($B/b E/-$)	7 ($-/b e/e$)	

58. a. Foi detectada uma série de alelos múltiplos: superduplo > único > duplo.

b. Embora a explicação para a parte a não explique todos os cruzamentos, ela não leva em conta a esterilidade da fêmea nem a origem da planta superduplo de uma variedade de flores duplas.

60. a. Um cruzamento triíbrido daria uma proporção 63:1. Portanto, existem três loci *R* segregando nesse cruzamento.

b. P	$R_1/R_1 ; R_2/R_2 ; R_3/R_3 \times r_1/r_1 ; r_2/r_2 ; r_3/r_3$	
F_1	$R_1/r_1 ; R_2/r_2 ; R_3/r_3$	
F_2	27 $R_1/- ; R_2/- ; R_3/-$	vermelho
	9 $R_1/- ; R_2/- ; r_3/r_3$	vermelho
	9 $R_1/- ; r_2/r_2 ; R_3/-$	vermelho
	9 $r_1/r_1 ; R_2/- ; R_3/-$	vermelho
	3 $R_1/- ; r_2/r_2 ; r_3/r_3$	vermelho
	3 $r_1/r_1 ; R_2/- ; r_3/r_3$	vermelho
	3 $r_1/r_1 ; r_2/r_2 ; R_3/-$	vermelho
	1 $r_1/r_1 ; r_2/r_2 ; r_3/r_3$	branco

c. (1) Para obter uma proporção 1:1, apenas um dos genes pode ser heterozigoto. Um cruzamento representativo é $R_1/r_1 ; r_2/r_2 ; r_3/r_3 \times r_1/r_1 ; r_2/r_2 ; r_3/r_3$.

(2) Para obter uma proporção de 3 vermelhos : 1 branco, dois alelos devem estar se segregando e não podem estar no mesmo gene. Um cruzamento representativo é $R_1/r_1 ; R_2/r_2 ; r_3/r_3 \times r_1/r_1 ; r_2/r_2 ; r_3/r_3$.

(3) Para obter uma proporção 7 vermelhos : 1 branco, três alelos devem estar se segregando e não podem estar no mesmo gene. O cruzamento é $R_1/r_1 ; R_2/r_2 ; R_3/r_3 \times r_1/r_1 ; r_2/r_2 ; r_3/r_3$.

d. A fórmula $1 - (1/2)^n$ onde n = número de loci que estão segregando nos cruzamentos representativos na parte c.

64. a. e b. A epistasia está implicada, e o genótipo recessivo homozigoto branco parece bloquear a produção de cor por um segundo gene.

Suponha as seguintes relações de dominância: vermelho > laranja > amarelo. Designemos os alelos como:

vermelho	A^R
laranja	A^O
amarelo	A^Y

Os cruzamentos de 1 a 3 agora se tornam:

P	$A^O/A^O \times A^Y/A^Y$	$A^R/A^R \times A^O/A^O$	$A^R/A^R \times A^Y/A^Y$
F_1	A^O/A^Y	A^R/A^O	A^R/A^Y
F_2	3 $A^O/-$: 1 A^Y/A^Y	3 $A^R/-$: 1 A^O/A^O	3 $A^R/-$: 1 A^Y/A^Y

Cruzamento 4: Para fazer este cruzamento, você deve adicionar um segundo gene. Você também deve reescrever os cruzamentos 1 a 3 para incluir o segundo gene. Façamos *B* representar a expressão da cor e *b* bloquear sua expressão, produzindo branco. Os primeiros três cruzamentos tornam-se:

P	$A^O/A^O ; B/B \times A^Y/A^Y ; B/B$
	$A^R/A^R ; B/B \times A^O/A^O ; B/B$
	$A^R/A^R ; B/B \times A^Y/A^Y ; B/B$
F_1	$A^O/A^Y ; B/B$
	$A^R/A^O ; B/B$
	$A^R/A^Y ; B/B$
F_2	3 $A^O/- ; B/B$: 1 $A^Y/A^Y ; B/B$
	3 $A^R/- ; B/B$: 1 $A^O/A^O ; B/B$
	3 $A^R/- ; B/B$: 1 $A^Y/A^Y ; B/B$

O quarto cruzamento é

P	$A^R/A^R ; B/B \times A^R/A^R ; b/b$
F_1	$A^R/A^R ; B/b$
F_2	3 $A^R/A^R ; B/-$: 1 $A^R/A^R ; b/b$

Cruzamento 5: Para fazer este cruzamento, note que não há laranja. Portanto, os dois genitores devem possuir os alelos para vermelho e amarelo, e a expressão de vermelho deve ser bloqueada.

P	$A^Y/A^Y ; B/B \times A^R/A^R ; b/b$
F_1	$A^R/A^Y ; B/b$
F_2	9 $A^R/- ; B/-$: 3 $A^R/- ; b/b$: 3 $A^Y/A^Y ; B/-$: 1 $A^Y/A^Y ; b/b$
	vermelho branco amarelo branco

Cruzamento 6: Este cruzamento é idêntico ao cruzamento 5, exceto que laranja substitui amarelo.

P	$A^O/A^O ; B/B \times A^R/A^R ; b/b$
F_1	$A^R/A^O ; B/b$
F_2	9 $A^R/- ; B/-$: 3 $A^R/- ; b/b$: 3 $A^O/A^O ; B/-$: 1 $A^O/A^O ; b/b$
	vermelho branco laranja branco

Cruzamento 7: Neste cruzamento, amarelo é suprimido por *b/b*.

P	$A^R/A^R ; B/B \times A^Y/A^Y ; b/b$
F_1	$A^R/A^Y ; B/b$
F_2	9 $A^R/- ; B/-$: 3 $A^R/- ; b/b$
	vermelho branco

3 A^Y/A^Y ; $B/-$ amarelo
1 A^Y/A^Y ; b/b branco

66. a. Entrecruzamento de linhagens mutantes, todas com um fenótipo recessivo comum, é a base do teste de complementação. Esse teste é destinado a identificar o número de genes diferentes que podem mutar para um fenótipo particular. Neste problema, se a prole de determinado cruzamento ainda expressa o fenótipo serpear, as mutações não se complementam e são consideradas alelos do mesmo gene; se a prole é tipo selvagem, as mutações se complementam e as duas linhagens levam alelos mutantes de genes separados.

b. Esses dados identificam cinco grupos de complementação (genes).

c. mutante 1: $a^1/a^1 \cdot b^+/b^+ \cdot c^+/c^+ \cdot d^+/d^+ \cdot e^+/e^+$ (embora apenas os alelos mutantes geralmente sejam listados)

mutante 2: $a^+/a^+ \cdot b^2/b^2 \cdot c^+/c^+ \cdot d^+/d^+ \cdot e^+/e^+$

mutante 5: $a^5/a^5 \cdot b^+/b^+ \cdot c^+/c^+ \cdot d^+/d^+ \cdot e^+/e^+$

1/5 híbridos: $a^1/a^5 \cdot b^+/b^+ \cdot c^+/c^+ \cdot d^+/d^+ \cdot e^+/e^+$

fenótipo: não serpear

1 e 5 são ambos mutantes do gene A

(o cruzamento relevante: $a^+/a^+ \cdot b^2/b^2 \times a^2/a^2 \cdot b^5/b^5$ dá)

2/5 híbridos: $a^+/a^5 \cdot b^+/b^2 \cdot c^+/c^+ \cdot d^+/d^+ \cdot e^+/e^+$

fenótipo: tipo selvagem

2 e 5 são mutantes de genes diferentes

Capítulo 7

1. A dupla hélice de DNA é mantida unida por dois tipos de ligações: covalentes e de hidrogênio. As ligações covalentes são encontradas dentro de cada filamento linear e fortemente ligadas às bases, açúcares e grupos fosfato (ambos dentro de cada componente e entre componentes). As pontes de hidrogênio são encontradas entre os dois filamentos; uma ponte de hidrogênio forma-se entre uma base em um filamento e uma base no outro filamento no pareamento complementar. Essas pontes de hidrogênio são individualmente fracas, mas coletivamente bem fortes.

4. As helicases são enzimas que rompem as pontes de hidrogênio que mantêm juntos os filamentos de DNA em uma dupla hélice. Essa quebra é necessária para a síntese de RNA e de DNA. As topoisomerasas são enzimas que criam e relaxam a super-helicoidização da dupla hélice de DNA. A super-hélice em si é um resultado da torção da hélice de DNA quando os dois filamentos se separam.

6. Não. A informação do DNA depende de um mecanismo fiel de cópia. As regras estritas de complementariedade garantem que a replicação e a transcrição possam ser reproduzidas.

9. O cromossomo seria irrecuperavelmente fragmentado.

12. b. O RNA mais provavelmente conteria erros.

16. Se o DNA é bifilamentar, $A = T$, $G = C$ e $A + T + C + G = 100\%$. Se $T = 15\%$, então $C = [100 - 15(2)]/2 = 35\%$.

17. Se o DNA é bifilamentar, $G = C = 24\%$ e $A = T = 26\%$.

20. Sim. A replicação do DNA também é semiconservativa nos eucariontes diplóides.

22. 3'....GGAATTCTGATTGATGAATGACCCTAG....5'

26. Sem telomerase funcional, os telômeros se encurtariam a cada ciclo de replicação, levando a uma perda de informação codificantes essencial e morte. De fato, algumas observações atuais indicam que o declínio ou perda da atividade de telomerase têm um papel no mecanismo de envelhecimento em humanos.

28. As regras de Chargaff são que $A = T$ e $G = C$. Como essas igualdades não são observadas, a interpretação mais provável é que o DNA é unifilamentar. O fago primeiro teria que sintetizar um filamento complementar antes que pudesse começar a multiplicar cópias de si mesmo.

Capítulo 8

2. Em procariontes, a tradução começa na ponta 5' enquanto a ponta 3' ainda está sendo sintetizada. Em eucariontes, o processamento (revestimento, recomposição) está ocorrendo do lado 5' enquanto a ponta 3' ainda está sendo sintetizada.

8. Sim. Tanto a replicação quanto a transcrição são feitas por grandes máquinas moleculares com várias subunidades (o replissomo e a RNA polimerase II, respectivamente) e ambas requerem atividade de helicase na

forquilha da bolha. Entretanto, a transcrição só ocorre em um sentido e apenas um filamento de DNA é copiado.

10. a. A sequência original representa as sequências de consenso -35 e -10 (com o número correto de espaços intercalares) e um promotor bacteriano. O fator σ , como parte da RNA polimerase holoenzima, reconhece essas sequências e liga-se a elas.

b. As sequências mutadas (transpostas) não serão um sítio de ligação para o fator σ . A orientação de duas regiões com relação uma a outra não é correta; portanto, não será reconhecida como um promotor.

15. As auto-remoções dos íntrons são capazes de removê-los de um transcrito primário sem a necessidade de enzimas adicionais ou fonte de energia. Elas são um dos muitos exemplos de moléculas de RNA que são catalíticas, e, por essa propriedade, também são conhecidas como ribozimas. Com essa função adicional, o RNA é a única molécula biológica conhecida a codificar a informação genética e catalisar reações biológicas. Em termos simples, a vida possivelmente começou com uma molécula de RNA ou grupo de moléculas que desenvolveram a habilidade de se auto-replicar.

19. O RNA bifilamentar, composto de um filamento com sentido e um filamento anti-sentido complementar, pode ser usado em *C. elegans* (e como todos os organismos) para evitar seletivamente a síntese dos produtos gênicos codificados (uma descoberta que recebeu o prêmio Nobel de Medicina de 2006). Esse processo, chamado de silenciamento gênico, bloqueia a síntese da proteína codificada pelo gene endógeno, e é, portanto, equivalente ao nocaute do gene. Para testar se um mRNA específico codifica uma proteína embrionária essencial, injete o RNA bifilamentar produzido pelo mRNA em zigotos de embriões muito iniciais, ativando assim a via de RNAi. Os efeitos do nocaute do produto gênico específico podem então ser acompanhados observando-se o que ocorre nesses embriões comparado com os controles. Se a proteína codificada é essencial, o desenvolvimento embrionário deve ser perturbado quando seu gene é silenciado.

Capítulo 9

2. a. e b. 5' UUG GGA AGC 3'

c. e d. Com a suposição de que a matriz de leitura começa na primeira base,

$\text{NH}_3 - \text{Leu} - \text{Gli} - \text{Ser} - \text{COOH}$

Para o filamento de baixo, o mRNA é 5' GCU UCC CAA 3' e, com a suposição de que a matriz de leitura começa na primeira base, a cadeia de aminoácidos correspondente é

$\text{NH}_3 - \text{Ala} - \text{Ser} - \text{Gln} - \text{COOH}$.

6. Existem três códons para isoleucina: 5' AUU 3', 5' AUC 3' e 5' AUA 3'. Os anticódons possíveis são 3' UAA 5' (complementar), 3' UAG 5' (complementar) e 3' UAI 5' (oscilante). Embora complementar, 5' UAU 3' também faria par de bases com 5' AUG 3' (metionina) devido à oscilação e, portanto, não seria uma alternativa aceitável.

11. A estrutura quaternária é devida a interações de subunidades de uma proteína. Neste exemplo, a atividade enzimática que está sendo estudada pode ser a de uma proteína que consiste em duas subunidades diferentes. Os polipeptídeos das subunidades são codificados por genes separados e não-ligados.

15. Não. A enzima pode precisar de modificações pós-traducionais para ser ativa. As mutações nas enzimas necessárias para essas modificações não seriam mapeadas no gene de isocitrato liase.

18. Com a suposição de que todas as três mutações do gene *P* são mutações sem sentido, três códons diferentes possíveis de fim (âmbar, ocre e opala) poderiam ser a causa. Uma mutação supressora seria específica para um tipo de códon sem sentido. Por exemplo, os supressores âmbar suprimiriam os mutantes âmbar, mas não opala ou ocre.

22. Mudanças únicas de aminoácidos podem resultar em mudanças no dobramento de proteínas, direcionamento de proteínas ou modificações pós-traducionais. Qualquer uma dessas mudanças poderia dar os resultados indicados.

29. Se o anticódon de uma molécula de tRNA foi alterado por mutação para ser de quatro bases de tamanho, com a quarta base do lado 5' do anticódon, isso iria suprimir a inserção. As alterações no ribossomo também podem induzir a mudança de matriz de leitura.

30. f, d, j, e, c, i, b, h, a, g.

Capítulo 10

2. Os mutantes O^C são alterações na sequência de DNA do operador que impedem a ligação do repressor *lac*. Como um operador controla apenas os genes no mesmo filamento de DNA, ele está em cis (no mesmo filamento).

5. Um gene é desligado ou inativado pelo “modulador” (geralmente chamado de *repressor*) no controle negativo, e o repressor deve ser removido para que ocorra a transcrição. Um gene é ligado pelo “modulador” (geralmente chamado de um *ativador*) no controle positivo, e o ativador deve ser adicionado ou convertido em uma forma ativa para que ocorra a transcrição.

10. Se um óperon estiver governando ambos os genes, então uma mutação de mudança de matriz pode fazer com que o códon de fim que separa os dois genes seja lido como um códon com sentido. Portanto, o segundo produto gênico seria incorreto para todos os aminoácidos. Entretanto, não existem óperons multigênicos conhecidos nos eucariontes. A alternativa e melhor explicação é que ambas as funções enzimáticas sejam feitas por partes diferentes do mesmo produto gênico. Nesse caso, uma mudança de matriz de leitura além da primeira função, carbamil fosfato sintetase, resultará na segunda metade da molécula de proteína sendo não-funcional.

13. A mutação S é uma alteração em *lacI* tal que a proteína repressora liga-se ao operador, independentemente de o indutor estar presente. Em outras palavras, é uma mutação que inativa o sítio alostérico que se liga a um indutor, mas não afeta a habilidade do repressor em se ligar ao sítio operador. A dominância da mutação S é devida à ligação do repressor mutante, mesmo nas circunstâncias em que o repressor normal não se liga ao DNA (na presença de um indutor). As mutações reversas constitutivas que estão mapeadas em *lacI* são eventos mutacionais que inativam a habilidade desse repressor em se ligar ao operador. As mutações constitutivas reversas que estão mapeadas no operador alteram a sequência do DNA operador de tal modo que não permitirá a ligação a nenhuma molécula repressora (tipo selvagem ou repressor mutante).

16. As mutações em *CI*, *cII* e *cIII* afetariam todas a lisogenia: *cI* codifica o repressor, *cII* codifica um ativador de P_{RE} e *cIII* codifica uma proteína que protege *cII* da degradação. As mutações em *N* (um antiterminalizador) também afetariam a lisogenia porque sua função é necessária para a transcrição dos genes *cII* e *cIII*, mas também é necessária para os genes que têm papel na lise. As mutações no gene que codifica a integrase (*int*) também afetariam a habilidade de um fago mutante em sofrer lisogenia.

Capítulo 11

2. Em geral, o estado básico de um gene bacteriano é “ligado”. Assim, o início da transcrição é evitado ou reduzido se a ligação da RNA polimerase é bloqueada. Em contraste, o estado básico nos eucariontes é “desligado”. Assim, a maquinaria transcricional (incluindo a RNA polimerase II e os fatores gerais de transcrição associados) não pode ligar-se ao promotor na ausência de outras proteínas reguladoras.

6. Entre as mutações que podem evitar que uma linhagem de levedura mude de tipo reprodutor estariam as mutações nos genes *HO* e *HMRa*. O gene *HO* codifica uma endonuclease que corta o DNA para iniciar a mudança, e o locus *HMRa* contém o “cassete” de informação genética não-expressa para o tipo reprodutivo MATa.

9. O termo *herança epigenética* é usado para descrever alterações herdáveis nas quais a própria sequência de DNA não é mudada. Ela pode ser definida operacionalmente como a herança dos estados de cromatina de uma geração celular para a seguinte. Imprinting genômico, inativação do cromossomo X e variação de efeito de posição são alguns desses exemplos.

13. A herança da estrutura da cromatina é tida como responsável pela herança da informação epigenética. Essa herança é devida à herança do código de histonas e também pode incluir a herança de padrões de metilação do DNA.

16. a. De D a J; o transcrito primário incluirá todos os éxons e íntrons.

b. E, G e I; todos os íntrons serão removidos.

c. A, C e L; as regiões do promotor e acentuador irão ligar-se a vários fatores de transcrição que podem interagir com a RNA polimerase.

20. Um gene não-expresso devido a alteração em sua sequência de DNA nunca será expresso e será herdado de geração a geração. Um gene epigeneticamente inativado ainda pode ser regulado. A estrutura da cromatina pode mudar no curso do ciclo celular, como, por exemplo, quando os fatores de transcrição modificam o código de histonas.

22. A estrutura da cromatina afeta muito a expressão gênica. Os transgenes inseridos em regiões de eucromatina mais provavelmente serão capazes de expressão do que os inseridos em regiões de heterocromatina.

Capítulo 12

2. O gene primário de regra dos pares *eve* (*even-skipped*) seria expresso em sete faixas ao longo do eixo A-P do blastoderma avançado.

6. Se você diagramar esses resultados, verá que a deleção de um gene que funciona posteriormente permite que os segmentos mais anteriores se estendam em uma direção posterior. A deleção de um gene anterior não permite a extensão do segmento próximo mais posterior em uma direção anterior. Os genes *gap* ativam *Ubx* tanto no segmento torácico como no abdominal, enquanto os genes *abd-A* e *Abd-B* são ativados apenas nos segmentos abdominais médio e posterior. O funcionamento dos genes *abd-A* e *Abd-B* nesses segmentos de algum modo impede a expressão de *Ubx*. Entretanto, se *abd-A* e *Abd-B* forem deletados, *Ubx* pode ser expresso nessas regiões.

9. a. Um gene de regra dos pares.

b. Veja a expressão do mRNA do gene candidato em um padrão repetido de sete faixas ao longo do eixo A-P do embrião em desenvolvimento.

c. Não. Um embrião mutante para o gene *gap* *Krüppel* teria falta de muitos segmentos anteriores. Esse efeito seria epistático à expressão de um gene de regra dos pares.

12. a. O homeodomínio é um domínio de proteína conservado contendo 60 aminoácidos encontrados em um número significativo de fatores de transcrição. Qualquer proteína que contenha um homeodomínio funcional é quase certamente um fator de transcrição de ligação a uma sequência específica do DNA.

b. O gene *eyeless* (assim denominado para um fenótipo mutante) regula o desenvolvimento do olho em *Drosophila*. Você esperaria que ele fosse expresso apenas nas células que dão origem aos olhos. Para testar essa previsão, a visualização do local da expressão do mRNA de *eyeless* por hibridização *in situ* e da proteína *Eyeless* por métodos imunológicos deve ser feita. Por manipulação genética, o gene *eyeless* pode ser expresso em tecidos nos quais ele não é geralmente expresso. Por exemplo, quando *eyeless* é ligado em células destinadas a formar pernas, formam-se olhos nas pernas.

c. Os experimentos transgênicos mostraram que o gene *Small eye* de camundongos e o gene *eyeless* de *Drosophila* são tão similares que o gene de camundongo pode substituir *eyeless* quando introduzido em *Drosophila*. Como na resposta da parte b, quando o gene *Small eye* é expresso em *Drosophila*, mesmo nas células destinadas a formar pernas, formam-se olhos nas pernas. (Entretanto, os “olhos” não são olhos de camundongo, pois *Small eye* e *eyeless* atuam como comandos master que ligam toda a cascata de genes necessários para fazer o olho — neste caso, um olho de *Drosophila*.)

16. A proteína GLP-1 está situada nas duas células anteriores do embrião de quatro células de *C. elegans* por repressão de sua tradução nas duas células posteriores. A repressão da tradução de GLP-1 requer o controle espacial da região 3' UTR (SCR). A deleção de SCR irá permitir a expressão de *glp-1* em ambas as células, anterior e posterior. Em ambos os mutantes, heterozigotos e homozigotos, você esperaria a expressão da proteína GLP-1 em todas as células.

Capítulo 13

2. Como as bactérias têm genomas pequenos (cerca de 3 pares de Mb) e essencialmente sem sequências repetidas, o enfoque shotgun de genoma inteiro seria usado.

4. Um arcabouço também é chamado de super-hélice. Um contig é uma sequência de leituras superpostas montadas em uma unidade, e um arcabouço é uma coleção de contigs unidos.

9. Sim. O operador é o local no qual o repressor se liga funcionalmente pelas interações de sequências de DNA e da proteína repressora.

14. Você pode determinar se o clone de cDNA é um monstro ou não pelo alinhamento da sequência de cDNA com a sequência genômica. (Estão disponíveis programas de computação para fazer tais alinhamentos.) A sequência é derivada de dois sítios diferentes? O cDNA mapeia dentro de uma região (do tamanho do gene) no genoma ou em duas regiões diferentes? Os íntrons podem complicar a questão.

15. a. Como o código triplo é redundante, mudanças na sequência de nucleotídeos do DNA (especialmente os nucleotídeos da terceira posição de um códon) podem ocorrer sem mudar sua proteína codificada.

b. As sequências de proteínas podem ser expressas para evoluir e divergir mais lentamente do que os genes que as codificam.

20. A montagem correta de regiões grandes e quase idênticas é problemática com ambos os métodos de sequenciamento genômico. Entretanto, o método shotgun é menos efetivo em encontrar essas regiões do que o baseado em clone. Este último método tem a vantagem adicional de facilitar o acesso ao clone suspeito para posterior análise.

23. 15% são funções gênicas essenciais (tais como enzimas necessárias para a replicação do DNA ou síntese de proteínas).

25% são auxotróficos (enzimas necessárias para a síntese de aminoácidos ou metabolismo de açúcares etc.).

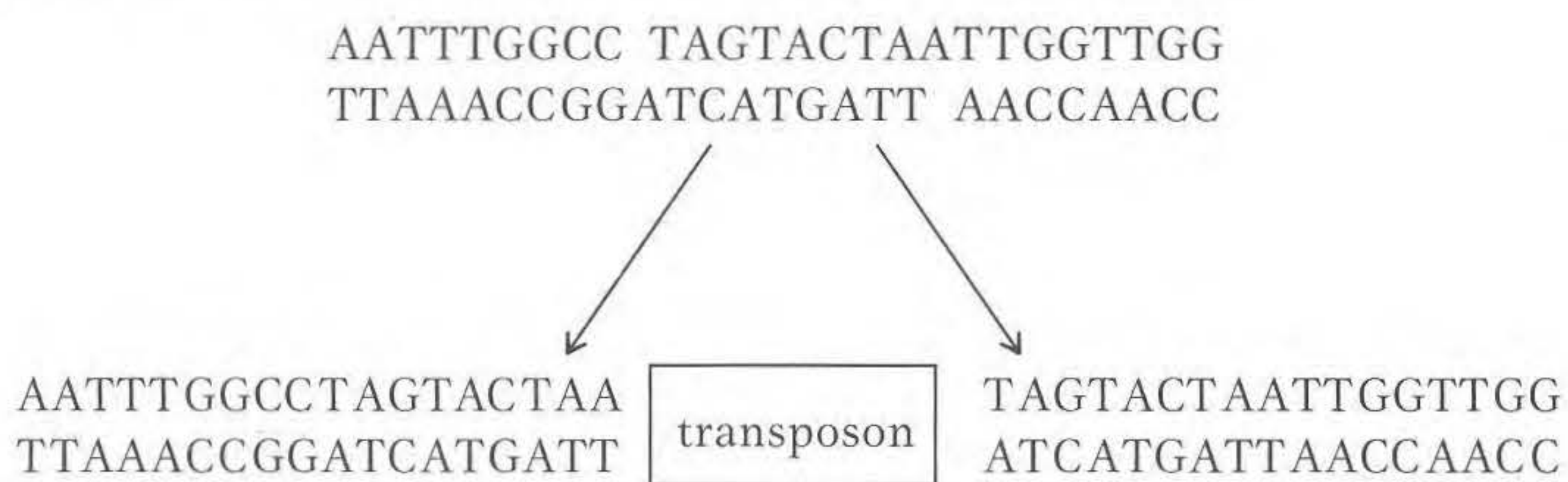
60% são redundantes ou vias não-testadas (genes para histonas, tubulina, RNA ribossômicos etc. estão presentes em várias cópias, a levedura pode precisar de muitos genes apenas sob situações únicas ou especiais, ou de outros modos que não são necessários à vida no laboratório).

Capítulo 14

2. Boeke, Fink e seus colaboradores demonstraram que a transposição do elemento *Ty* em leveduras é por um RNA intermediário. Eles construíram um plasmídeo usando um elemento *Ty* no qual inseriram não só um promotor que pode ser ativado pela galactose, mas também um íntron na região codificante do elemento *Ty*. Primeiro, a frequência de transposição foi muito aumentada pela adição de galactose, indicando que um aumento na transcrição (e produção de RNA) estava correlacionado com as taxas de transposição. Mais importante é que, após a transposição, eles descobriram que o DNA *Ty* recém-transposto não tinha a sequência íntron. Como a remoção do íntron só ocorre durante o processamento do RNA, deve ter havido um RNA intermediário no evento de transposição.

6. Alguns elementos de transposição desenvolveram estratégias para inserir-se em abrigos de segurança, regiões do genoma onde sofrem danos mínimos. Os abrigos de segurança incluem genes duplicados (tais como genes de tRNA ou rRNA) e outros elementos de transposição. Os abrigos de segurança em genomas bacterianos devem ser sequências muito específicas entre genes ou genes de rRNA repetido.

9. O corte desencontrado levará a uma duplicação de sítio-alvo de nove pares de bases flanqueadores do transposon inserido.



12. Não seria surpreendente encontrar um elemento SINE em um íntron de um gene, em vez de um éxon. O processamento do pré-mRNA removeria o elemento de transposição como parte do íntron e a tradução da enzima FB não seria efetuada.

Capítulo 15

1. Você precisa conhecer a matriz de leitura da possível mensagem.

4. Com a suposição de substituições de um único par de bases, CGG pode ser mudado para CGU, CGA, CGG ou AGG e ainda codificar arginina.

9. A seguinte lista de observações concorda com “câncer é uma doença genética”:

- (1) Alguns cânceres são herdados como características mendelianas simples altamente penetrantes.
- (2) A maioria dos agentes carcinogênicos também são mutagênicos.
- (3) Vários oncogenes foram isolados de vírus tumorais.
- (4) Vários genes que levam à suscetibilidade a tipos particulares de câncer foram mapeados, isolados e estudados.
- (5) Oncogenes dominantes foram isolados de células tumorais.
- (6) Alguns cânceres estão altamente correlacionados a rearranjos cromossômicos específicos.

12. A T mal pareada seria corrigida para C, e o ACG resultante, após a transcrição, seria 5' UGC 3' e codificaria cisteína. Ou, se o outro filamento fosse corrigido, ATG seria transcrito em 5' UAC 3' e codificaria tirosina.

17. Se uma quebra bifilamentar acidental ocorrer após a replicação, ela pode ser reparada pela SDSA (helicoidização de filamento dependente de síntese). Nesse processo, o Rad51-filamento de DNA “procura” a cromátide-irmã não-danificada quanto à sequência complementar que será usada

como molde para reparo livre de erro. [Se ocorrer uma quebra bifilamentar antes da síntese, então pode ser usado o mecanismo de reparo NHEJ (junção de pontas não-homólogas) livre de erro.] Na meiose, Rad51 associa-se à proteína Dmc1 específica da meiose. O filamento Rad51-Dmc1-filamento de DNA agora é capaz de “encontrar” sua sequência complementar no cromossomo homólogo em vez de em sua cromátide-irmã. Adicionalmente, o “reparo” de quebras bifilamentares geradas por SpoII na meiose pode resultar em recombinação.

22. Muitos sistemas de reparo estão disponíveis: reversão direta, reparo por excisão, reparo acoplado a transcrição e junção de pontas não-homólogas.

23. Sim, é mutagênica. Causará transições CG-TA.

28. a. Uma falta de revertentes sugere uma deleção ou uma inversão dentro do gene.

b. Para compreender esses dados, lembre que metade da prole deve vir do genitor tipo selvagem.

Prototrófico A: Como 100% da prole é de prototróficos, uma reversão no sítio original mutante deve ter ocorrido.

Prototrófico B: 50% da prole são prototróficos parentais, e os prototróficos restantes, 28%, são o resultado da nova mutação. Observe que os 28% são aproximadamente iguais aos 22% auxotróficos. A sugestão é que ocorreu uma mutação supressora não-ligada, produzindo segregação independente com o mutante *nic-2*.

Prototrófico C: Existem 496 prototróficos “revertentes” (os outros 500 são prototróficos parentais) e quatro auxotróficos. Isso sugere que ocorreu uma mutação supressora em um sítio muito próximo [$100\%(4 \times 2)/1.000 = 0,8$ u.m.] para a mutação original.

32. Xeroderma pigmentoso é um distúrbio genético heterogêneo e é causado por mutações em qualquer um dos vários genes que tomam parte no processo de NER (reparo por excisão de nucleotídeo). Como na descoberta de uma outra proteína na via NHEJ por pesquisa na parada da linhagem celular 2BN, esse paciente poderia ter uma mutação em um gene ainda não-conhecido que codifica uma proteína necessária para NER.

Capítulo 16

1. MM N OO seriam classificados como $2n - 1$; MM NN OO seriam classificados como $2n$; e MMM NN PP seriam classificados como $2n + 1$.

4. Haveria um possível quadrivalente.

7. Sete cromossomos.

9. As células destinadas a se tornarem grãos de pólen podem ser induzidas por tratamento com frio para crescer em embrióides. Esses embrióides podem então ser cultivados em ágar para formar plântulas monoplóides.

11. Sim.

14. Não.

16. Um fragmento acêntrico não pode ser alinhado ou movido na meiose (ou mitose) e é conseqüentemente perdido.

19. Deleções muito grandes tendem a ser letais, provavelmente devido a desequilíbrio genômico ou revelação de genes letais recessivos. Portanto, a alça de pareamento muito grande observada é, mais provavelmente, de uma inversão heterozigota.

21. A síndrome de Williams é o resultado de uma deleção da região 7q11.23 do cromossomo 7. A síndrome do *cri du chat* resulta de uma deleção de uma parte significativa do braço curto do cromossomo 5 (especificamente, as bandas 5p15.2 e 5p15.3). Tanto a síndrome de Turner (XO) quanto a de Down (trissomia do 21) resultam de não-disjunção meiótica. O termo “síndrome” é usado para descrever um conjunto de fenótipos (em geral complexos e variados) que estão em regra presentes juntos.

26. A ordem é *b a c e d f*.

Alelo	Banda
<i>b</i>	1
<i>a</i>	2
<i>c</i>	3
<i>e</i>	4
<i>d</i>	5
<i>f</i>	6

27. Os dados sugerem que um ou ambos os pontos de quebra da inversão estão situados dentro de um gene essencial, causando uma mutação letal recessiva.

30. a. Quando cruzado com fêmeas amarelas, os resultados seriam

X^e/Y^{e+}	machos cinza
X^e/X^e	fêmeas amarelas

b. Se o alelo e^+ fosse translocado para um autossomo, a prole seria a seguinte, onde "A" indica autossomo:

P	$A^{e+}/A ; X^e/Y \times A/A X^e/X^e$	
F_1	$A^{e+}/A ; X^e/X^e$	fêmea cinza
	$A^{e+}/A ; X^e/Y$	macho cinza
	$A/A ; X^e/X^e$	fêmea amarela
	$A/A ; X^e/Y$	macho amarelo

32. Síndrome de Klinefelter macho XXY
Síndrome de Down trissomia do 21
Síndrome de Turner fêmea XO

36. a. Se um hexaplóide fosse cruzado com um tetraplóide, o resultado seria pentaplóide.

b. Cruzar A/A com $a/a/a/a$ para obter $A/a/a$.

c. O modo mais fácil é expor as células da planta A/a^* a colchicina por uma divisão celular, o que resultaria na duplicação dos cromossomos para produzir $A/A/a^*/a^*$.

d. Cruzar um hexaplóide ($a/a/a/a/a/a$) com um diplóide (A/A) para obter $A/a/a/a$.

38. a. A proporção de plantas com folhas normais para folhas potato será 5:1.

b. Se o gene não está no cromossomo 6, deve haver uma proporção de 1:1 de plantas com folhas normais para folhas potato.

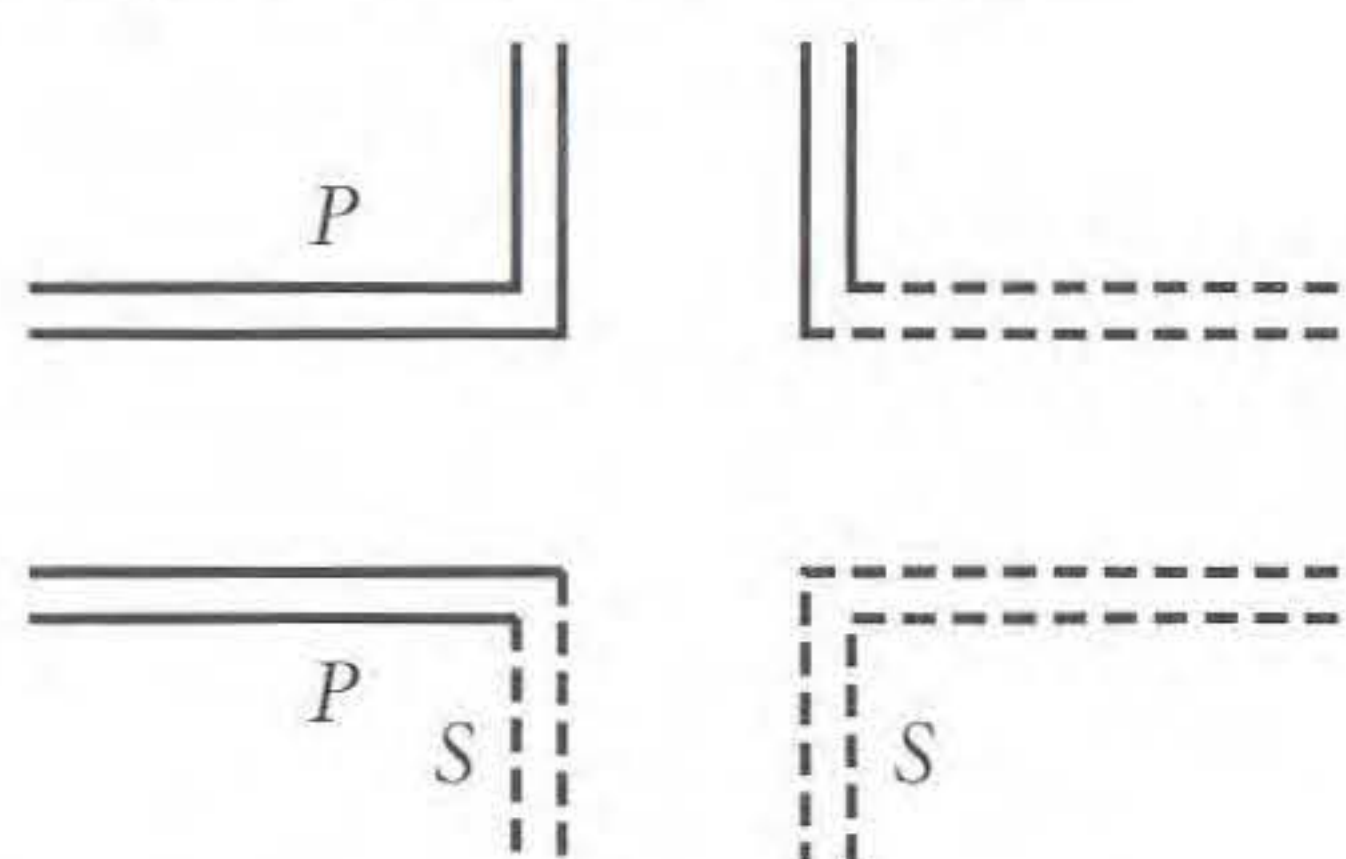
42. a. A planta aberrante é semi-estéril, o que sugere uma inversão. Como as frequências de recombinação $d-f$ e $y-p$ na planta aberrante são normais, a inversão deve implicar de b a x .

b. Para obter a prole de recombinação quando ocorreu uma inversão, é necessária a ocorrência ou de um duplo crossing dentro da região invertida, ou de crossings únicos entre f e a inversão, o que ocorreria em algum lugar entre f e y .

44. A planta original tem dois cromossomos com translocações recíprocas que colocam os genes P e S bem próximos. Devido à próxima ligação, uma proporção sugerindo um cruzamento monóibrido, em vez de díibrido, foi observada tanto com autofecundação quanto com um cruzamento-teste. Todos os gametas são férteis devido à homozigose.

planta original:	$P S/p s$
testador:	$p s/p s$

Prole F_1 : heterozigotos para a translocação:



O modo mais fácil de testar essa hipótese é observar os cromossomos dos heterozigotos na meiose I.

50. Os genitores originais devem ter tido a seguinte constituição cromossômica:

<i>G. hirsutum</i>	26 grandes, 26 pequenas
<i>G. thurberi</i>	26 pequenas
<i>G. herbaceum</i>	26 grandes

G. hirsutum é um derivado poliplóide de um cruzamento entre as duas espécies do Velho Mundo, o que poderia ser facilmente verificado observando-se os cromossomos.

52. a. Perda de um cromossomo X no feto em desenvolvimento após o estágio de duas células.

b. Não-disjunção levando à síndrome de Klinefelter (XXY), seguida de um evento de não-disjunção em uma célula para o cromossomo Y após o estágio de duas células, resultando em XX e XXY.

c. Não-disjunção do X no estágio de uma célula.

d. Fusão de zigotos XX e XY (de fertilizações separadas seja de dois zigotos ou de um zigoto e um glóbulo polar, um espermatozóide com X e um com Y).

e. Não-disjunção do X no estágio de duas células ou depois.

55. a. Cada mutante é cruzado com um tipo selvagem, ou

$$m \times m^+$$

As tétrades resultantes (óctades) mostram segregação 1:1, indicando que cada mutante é o resultado de uma mutação em um único gene.

b. Os resultados do cruzamento das duas linhagens mutantes indicam que ambas as linhagens são mutantes do mesmo gene:

$$m_1 \times m_2$$

ou que são mutantes em genes diferentes mas proximamente ligados:

$$m_1 m_2^+ \times m_1^+ m_2$$

c. e d. Como a prole fenotipicamente preta pode resultar de não-disjunção (note que, nos casos C e D, preto aparece em conjunto com esporos abortados), o mutante 1 e o mutante 2 são provavelmente mutantes em genes diferentes mas proximamente ligados. O cruzamento é portanto

$$m_1 m_2^+ \times m_1^+ m_2$$

O caso A é uma tétrade NPD e seria o resultado de um crossing duplo de quatro filamentos

$m_1^+ m_2^+$	preto
$m_1^+ m_2^+$	preto
$m_1 m_2$	castanho
$m_1 m_2$	castanho

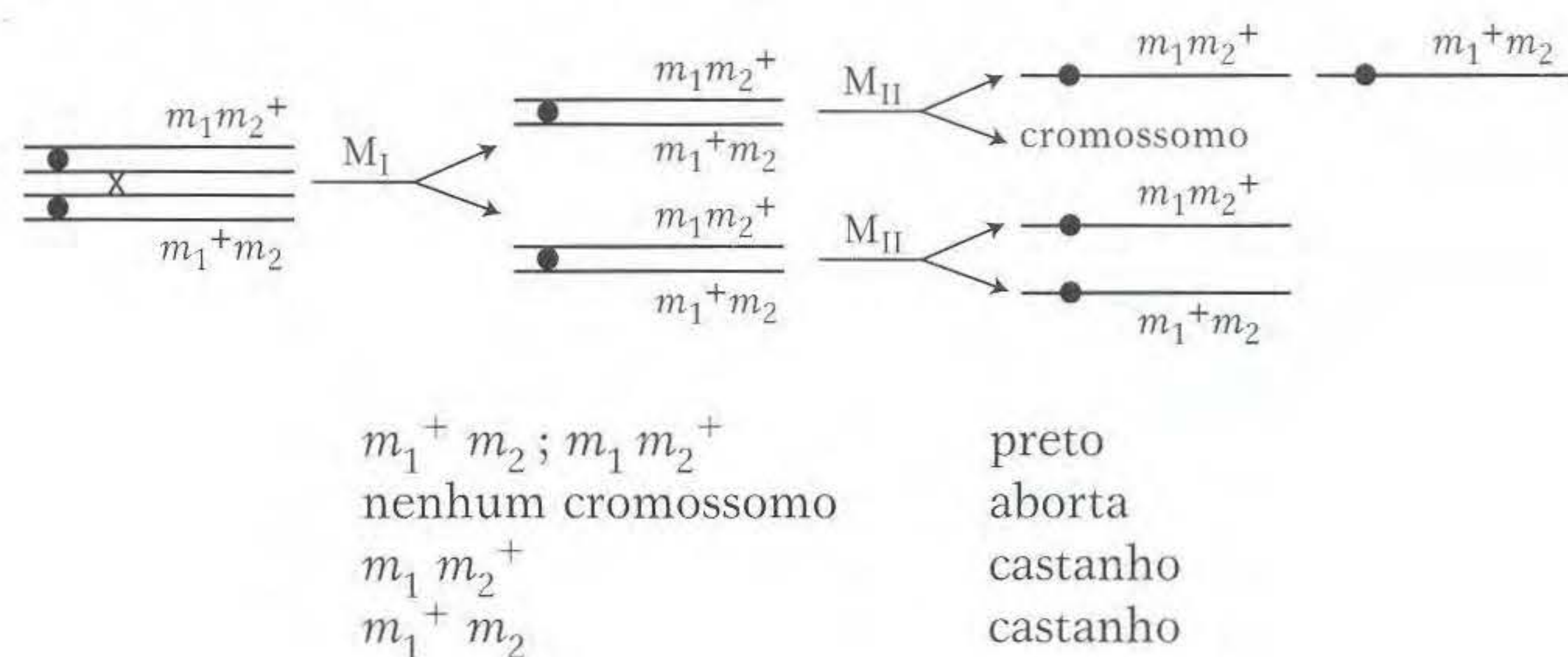
O caso B é um tetratipo e seria o resultado de um único crossing entre um dos genes e o centrômero.

$m_1^+ m_2^+$	preto
$m_1^+ m_2$	castanho
$m_1 m_2^+$	castanho
$m_1 m_2$	castanho

O caso C é o resultado de uma não-disjunção na meiose I.

$m_1^+ m_2^+ ; m_1^+ m_2^+$	preto
$m_1^+ m_2^+ ; m_1^+ m_2^+$	preto
nenhum cromossomo	aborta
nenhum cromossomo	aborta

O caso D resulta de recombinação entre um dos genes e o centrômero seguido de não-disjunção na meiose II. Por exemplo,



Capítulo 17

1. A frequência de um alelo em uma população pode ser alterada por seleção natural, mutação, migração, reprodução não-aleatória e deriva genética (erros de amostragem).

3. A frequência de B/B é $p^2 = 0,64$, e a frequência de B/b é $2pq = 0,32$.

5. a. $p' = 0,5[(0,5)(0,9) + 0,5(1,0)]/[(0,25)(0,9) + (0,5)(1,0) + (0,25)(0,7)] = 0,53$

b. 0,75

8. 0,65

10. a. A população está em equilíbrio.

b. A reprodução é aleatória com relação ao tipo sanguíneo.

12. a. A frequência na população é igual a $q^2 = 0,01$.

b. Existem 10 homens daltônicos para cada mulher daltônica (q/q^2).

c. 0,018.

d. Todas as crianças seriam fenotipicamente normais apenas se a mãe fosse homozigota para o alelo de visão normal ($p^2 = 0,81$). O genótipo do pai não tem importância e, portanto, poderia ser ignorado.

e. A frequência de mulheres daltônicas será 0,12 e de homens daltônicos será 0,2.

f. Da análise dos resultados na parte e, a frequência do alelo de daltonismo será 0,2 nos homens (a mesma que nas mulheres da geração anterior) e 0,4 nas mulheres.

15. Antes da migração, $q_A = 0,1$ e $q_B = 0,3$ nas duas populações. Como as duas populações são iguais em número, imediatamente após a migração, $q_{A+B} = \frac{1}{2}(q_A + q_B) = \frac{1}{2}(0,1 + 0,3) = 0,2$. No novo equilíbrio, a frequência de homens afetados é $q = 0,2$, e a frequência de mulheres afetadas é $q^2 = (0,2)^2 = 0,04$. (Daltonismo é uma característica ligada ao X.)

18. a. Muitos genes afetam o número de cerdas em *Drosophila*. A seleção artificial resultou em linhagens com os alelos para o maior número de cerdas. Algumas mutações podem ter ocorrido nas 20 gerações de cruzamentos seletivos, mas a maior parte da resposta é devida a alelos presentes na população original. A distribuição e recombinação geraram linhagens com alelos para maior número de cerdas.

b. A fixação de alguns alelos causando maior número de cerdas impede a reversão completa. Alguns alelos de alto número de cerdas não têm efeitos negativos na adaptabilidade; assim, nenhuma força está diminuindo o número de cerdas por causa desses loci.

c. A baixa fertilidade na linhagem com alto número de cerdas pode ser devida a pleiotropia ou ligação. Alguns alelos que causam alto número de cerdas também podem causar baixa fertilidade (pleiotropia). Os cromossomos com alelos para alto número de cerdas também podem levar alelos em loci diferentes que causam baixa fertilidade (ligação). Após a seleção artificial ter sido relaxada, os alelos de baixa fertilidade são selecionados contra por seleção natural. Algumas gerações de seleção relaxada permitem que os alelos de baixa fertilidade se recombinem, produzindo cromossomos de alto número de cerdas que não contêm alelos de baixa fertilidade. Quando a seleção é reaplicada, os alelos de baixa fertilidade são reduzidos em frequência ou separados dos loci de alto número de cerdas, e, assim, dessa vez há menos problema de fertilidade.

21. a. Custo genético = $sq^2 = 0,5(4,47 \times 10^{-3})^2 = 10^{-5}$

b. Custo genético = $sq^2 = 0,5(6,32 \times 10^{-3})^2 = 2 \times 10^{-5}$

c. Custo genético = $sq^2 = 0,3(5,77 \times 10^{-3})^2 = 10^{-5}$

Capítulo 18

1. Muitas características variam mais ou menos continuamente em uma ampla gama. Por exemplo, altura, peso, forma, cor, taxa reprodutiva, atividade metabólica etc. variam quantitativamente, e não qualitativamente. A variação contínua em geral é representada por uma curva em forma de sino, onde o fenótipo “médio” é mais comum que os extremos. A variação descontínua descreve os fenótipos facilmente classificáveis da genética mendeliana da forma das sementes, mutantes auxotróficos, anemia falciforme etc. Essas características mostram uma relação simples entre genótipo e fenótipo.

4. Desconhecidas são (1) as normas de reação para genótipos que afetam o QI, (2) a distribuição ambiental na qual os indivíduos se desenvolveram e (3) as distribuições genotípicas nas populações. Mesmo que elas fossem conhecidas, como a herdabilidade é específica para determinada população e seu ambiente, a diferença entre duas populações diferentes não pode receber um valor de herdabilidade.

6. a. $p(\text{homozigoto em 1 locus}) = 3(1/2)^3 = 3/8$

$p(\text{homozigoto em 2 loci}) = 3(1/2)^3 = 3/8$

$p(\text{homozigoto em 3 loci}) = 2(1/2)^3 = 2/8$

b. $p(0 \text{ letras maiúsculas}) = p(\text{todos os homozigotos recessivos}) = (1/4)^3 = 1/64$

$p(1 \text{ letra maiúscula}) = p(1 \text{ heterozigoto e 2 homozigotos recessivos}) = 3(1/2)(1/4)(1/4) = 3/32$

$p(2 \text{ letras maiúsculas}) = p(1 \text{ homozigoto dominante e 2 homozigotos recessivos})$

ou

$p(2 \text{ heterozigotos e 1 homozigoto recessivo}) = 3(1/4)^3 + 3(1/4)(1/2)^2 = 15/64$

$p(3 \text{ letras maiúsculas}) = p(\text{todos heterozigotos})$

ou

$p(1 \text{ homozigoto dominante, 1 heterozigoto e 1 homozigoto recessivo}) = (1/2)^3 + 6(1/4)(1/2)(1/4) = 10/32$

$p(4 \text{ letras maiúsculas}) = p(2 \text{ homozigotos dominantes e 1 homozigoto recessivo})$

ou

$p(1 \text{ homozigoto dominante e 2 heterozigotos}) = 3(1/4)^3 + 3(1/4)(1/2)^2 = 15/64$

$p(5 \text{ letras maiúsculas}) = p(2 \text{ homozigotos dominantes e 1 heterozigoto}) = 3(1/4)^2(1/2) = 3/32$

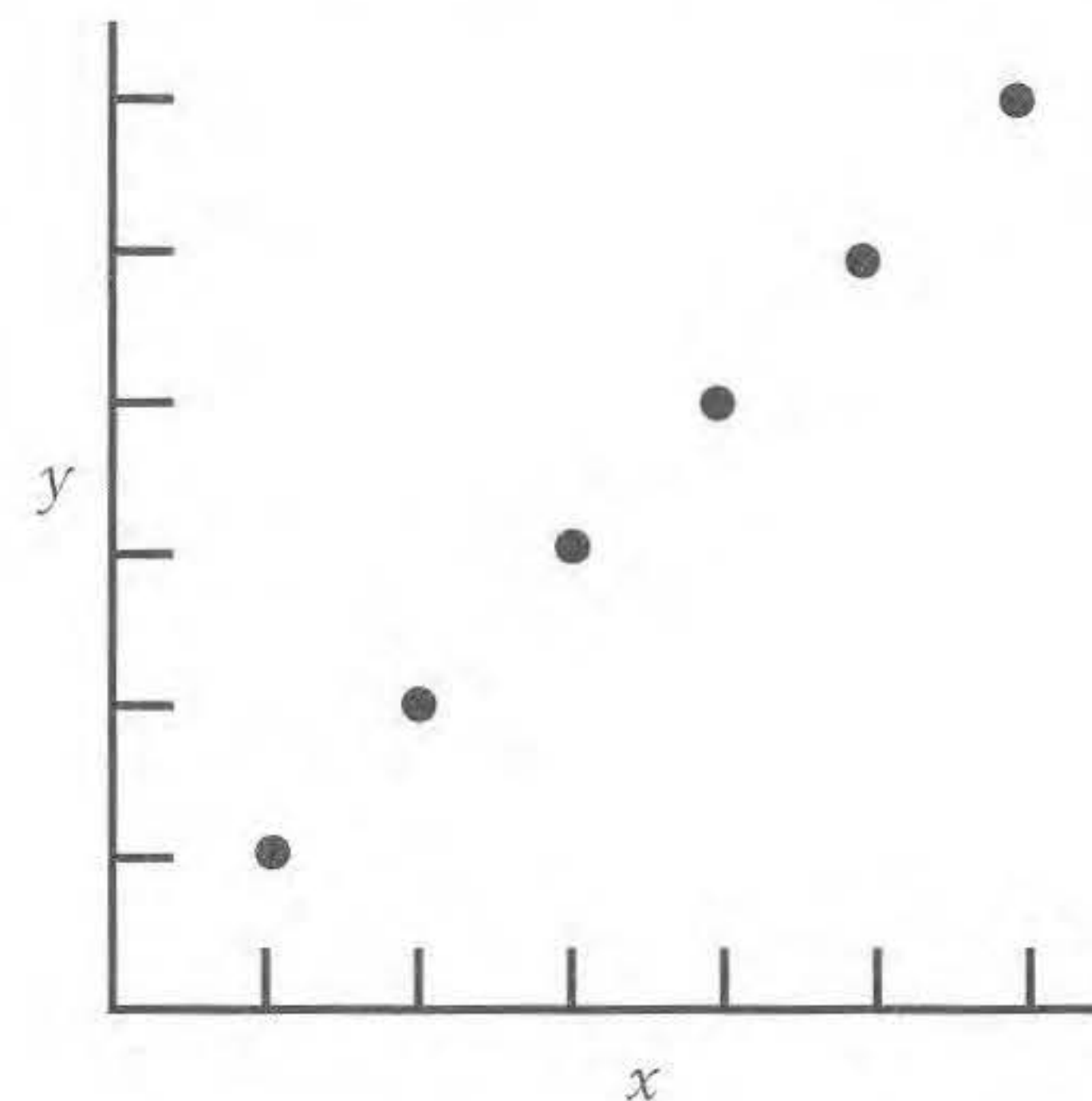
$p(6 \text{ letras maiúsculas}) = p(\text{todos os homozigotos dominantes}) = (1/4)^3 = 1/64$

8. A população descrita seria distribuída do seguinte modo:

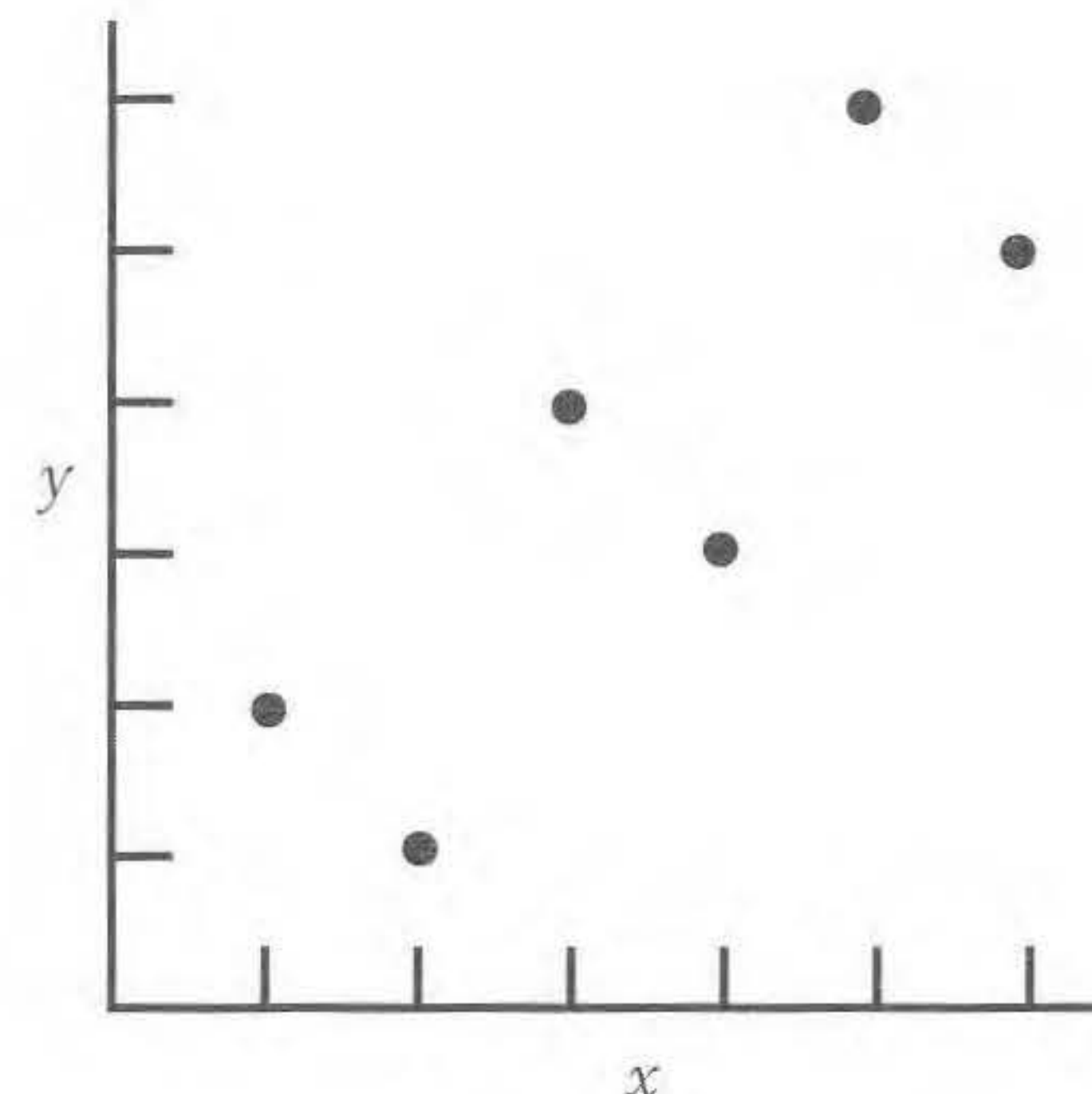
3 cerdas	19/64
2 cerdas	44/64
1 cerda	1/64

A classe de 3 cerdas conteria 7 genótipos diferentes, a classe de 2 cerdas conteria 19 genótipos diferentes e a classe de 1 cerda conteria apenas 1 genótipo. Seria muito difícil determinar a situação genética subjacente fazendo cruzamentos controlados e determinando as frequências da prole.

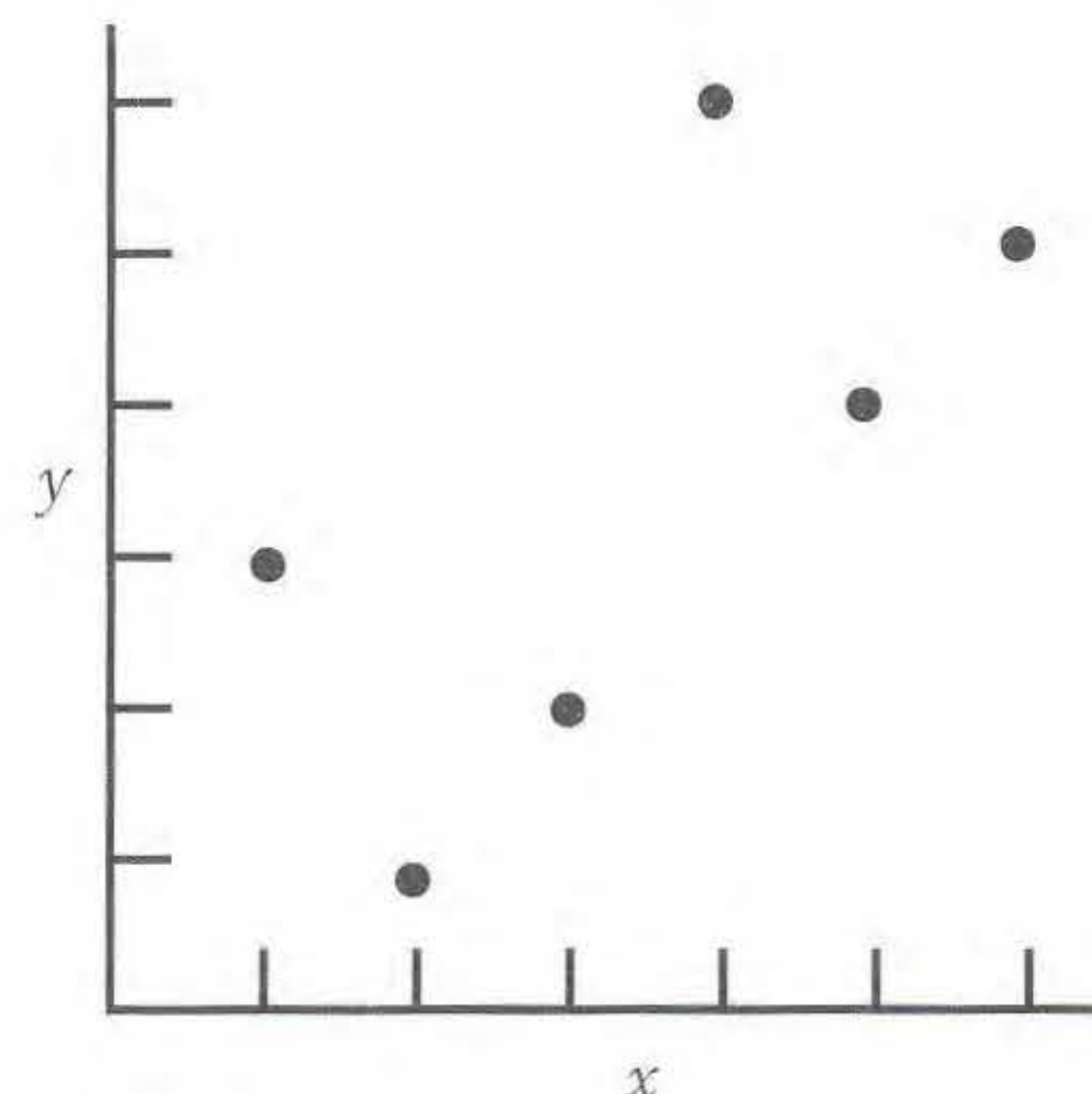
10. a.



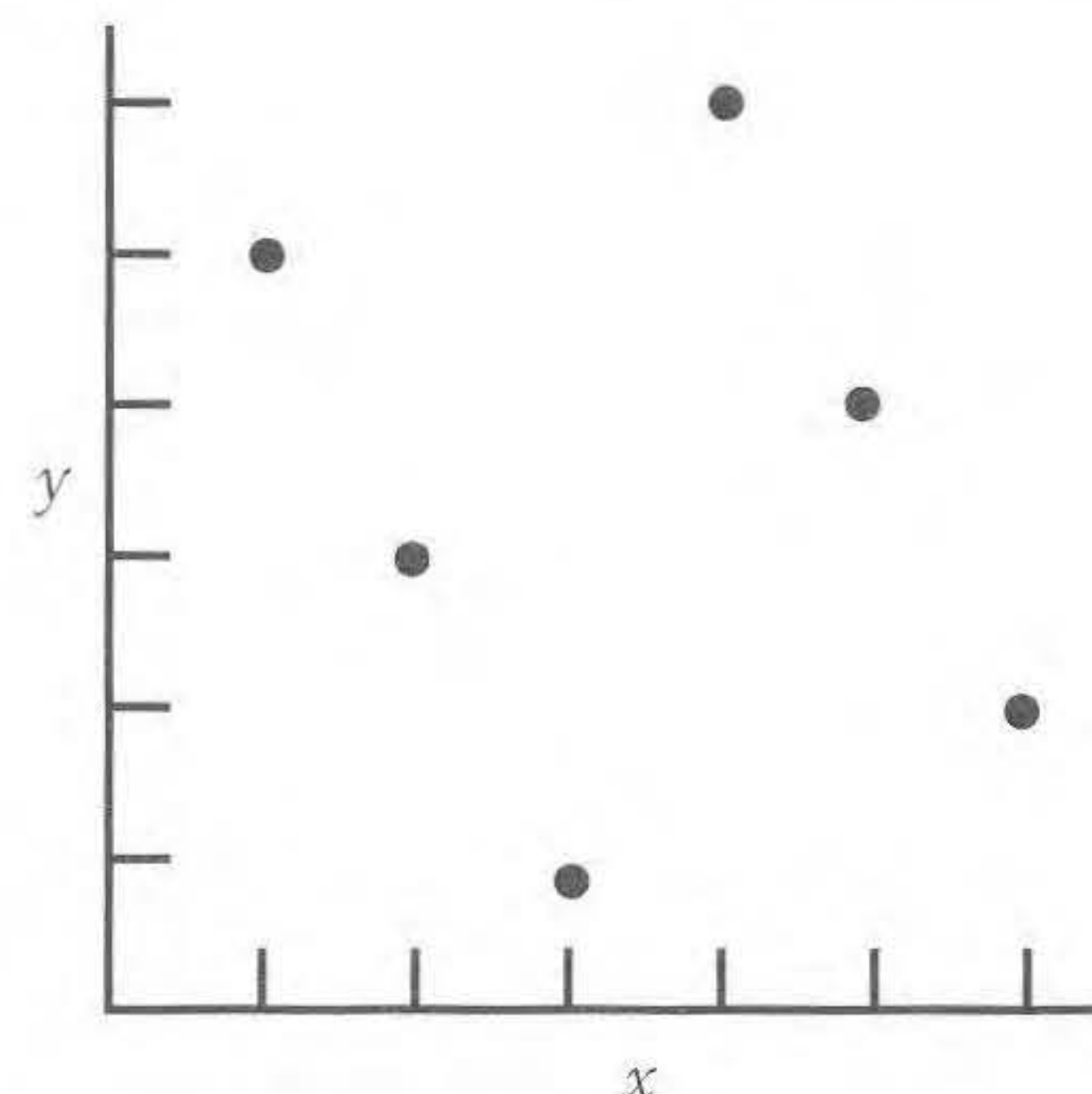
b.



c.



d.



a. $16[(1)(1) + (2)(2) + (3)(3) + (4)(4) + (5)(5) + (6)(6)] - (21/6)(21/6) = 2,92.$

$$\text{Desvio padrão } x = s_x = \sqrt{1/N \sum (x_i - \bar{x})^2} = 1,71$$

$$\text{Desvio padrão } y = s_y = \sqrt{1/N \sum (y_i - \bar{y})^2} = 1,71$$

Portanto, $r_{xy} = 2,92/(1,71)(1,71) = 1,0$. Os outros coeficientes de correlação são calculados de modo similar.

b. 0,83

c. 0,66

d. -0,20

14. a. A linha de regressão mostra a relação entre as duas variáveis. Ela tenta prever um (a altura do filho) pelo outro (a altura do pai). Se a relação for perfeitamente correlata, a inclinação da linha de regressão deve se aproximar de 1. Se você supõe que os indivíduos no extremo de qualquer espectro são homozigotos para os genes responsáveis por esses fenótipos, então sua prole mais provavelmente será heterozigota do que os indivíduos originais. Isto é, eles serão menos extremos. Adicionalmente, não há tentativa de incluir a contribuição materna para esse fenótipo.

b. Para os dados de Galton, regressão é uma estimativa de herdabilidade (h^2), supondo que existem poucas diferenças ambientais entre todos os pais e todos os filhos, tanto individualmente como em grupo. Entretanto, não é dada nenhuma evidência para determinar se as características são familiares mas não-herdáveis. Esses dados indicariam variação genética apenas se os parentes não compartilhassem ambientes comuns mais do que os não-aparentados fazem.

Capítulo 19

2. Os três princípios são: (1) organismos dentro de uma espécie variam de um para outro, (2) a variação é herdável e (3) diferentes tipos deixam números diferentes de proles em gerações futuras.

5. Uma população não se diferencia de outras populações por endogamia local se:

$$\mu \geq 1/N$$

e assim

$$N \geq 1/\mu$$

$$N \geq 10^5$$

7. Uma população não se diferenciará de outras populações por endogamia local se o número de indivíduos migrantes ≥ 1 por geração.

a. A migração não é suficiente para evitar endogamia local, e, assim, os resultados são aproximadamente os mesmos que no Problema 6.

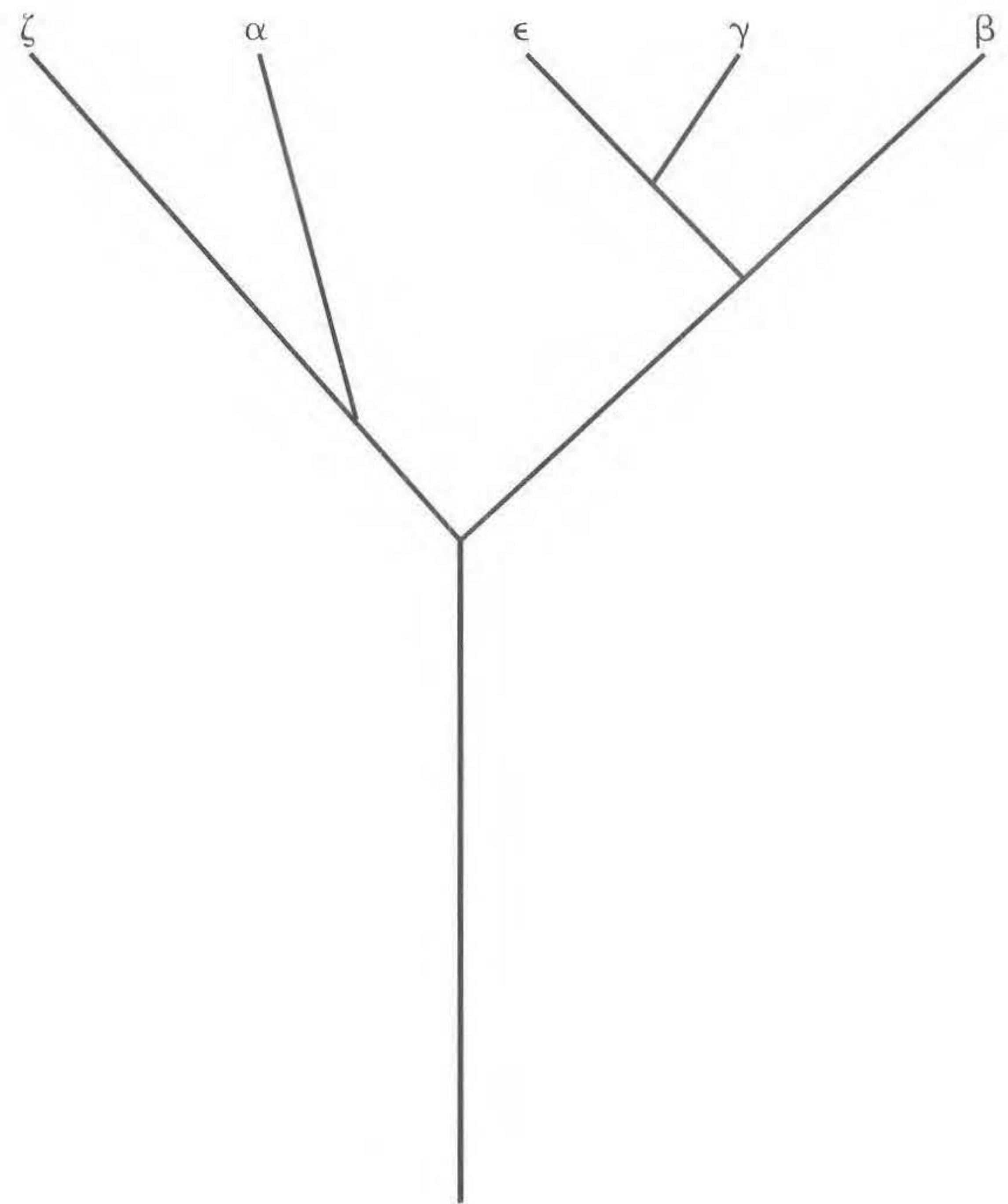
b. Há um imigrante por geração, e, assim, as populações não se diferenciarão e as frequências alélicas permanecerão as mesmas em todas as populações.

9. A adaptabilidade média para a população 1 [$p(A) = 0,5$, $p(B) = 0,5$] é 0,825. A adaptabilidade média para a população 2 [$p(A) = 0,1$, $p(B) = 0,1$] é 0,856. Existem quatro picos adaptativos: em $p(A) = 0,0$ ou 1,0 e em $p(B) = 0,0$ ou 1,0. (A adaptabilidade média será 0,90 em qualquer desses pontos.) Com a população 1, a direção de mudança tanto para $p(A)$ quanto para $p(B)$ será aleatória. Tanto as frequências mais altas quanto as mais baixas de ambos os alelos podem resultar em uma adaptabilidade média aumentada (embora existam algumas combinações que diminuiriam a adaptabilidade). Como a população 2 já está quase perto do pico adaptativo de $p(A) = 0,0$ e $p(B) = 0,0$, tanto $p(A)$ quanto $p(B)$ devem diminuir para aumentar a adaptabilidade média.

12. Sequências não-codificantes. Uma importante restrição da evolução gênica constitui os potenciais efeitos pleiotrópicos das mutações nas regiões codificantes. Esses efeitos podem ser evitados por mutações nas sequências regulatórias, que têm um papel importante na evolução da forma do corpo. As mudanças em sequências não-codificantes fornecem um mecanismo para alterar um aspecto da expressão gênica enquanto preserva o papel de proteínas pleiotrópicas em outros processos desenvolvimentais essenciais.

17. Todas as populações humanas têm alta frequência de i , I^A intermediária e I^B baixa. As variações que existem entre diferentes populações geográficas são mais provavelmente devidas à deriva genética. Não há evidência de que a seleção tenha algum papel quanto a esses alelos.

19.



Capítulo 20

2. A transcriptase reversa polimeriza o DNA com o uso de RNA como molde. Esse híbrido DNA-RNA é então tratado com hidróxido de sódio para degradar o RNA. (NaOH hidrolisa RNA catalisando a formação de um 2',3'-fosfato cíclico.)

4. A ligase é uma enzima essencial, dentro de todas as células, que fecha quebras no arcabouço açúcar-fosfato do DNA. Na replicação do DNA, a ligase une fragmentos de Okazaki para criar um filamento contínuo e, na clonagem, é usada para unir os vários fragmentos do DNA com o vetor. Se não for adicionada, o vetor e o DNA clonado simplesmente se separariam.

5. Cada ciclo leva 5 minutos e duplica o DNA. Em 1 hora, ocorreriam 12 ciclos; logo, o DNA seria amplificado $2^{12} = 4.096$ vezes.

8. Você poderia isolar o DNA da planta transgênica suspeita e sondar quanto à presença do transgene por hibridização de Southern.

12. a. O fenótipo transformado mapearia no mesmo locus. Se a substituição foi devida a um crossing duplo, as células transformadas não contêm DNA vetor. Se ocorreu um único crossing over, todo o vetor agora será parte do cromossomo linear de *Neurospora*.

b. O fenótipo transformado mapeará em um locus diferente do auxotrófico se o gene transformante for inserido ectopicamente (em outro local).

13. Tamanho, translocações entre cromossomos conhecidos e hibridização com sondas de locais conhecidos podem ser úteis na identificação de qual banda no gel de campo pulsado corresponde a determinado cromossomo.

19. A região do DNA que codifica tirosinase em DNA genômico de camundongo "normal" contém dois sítios de *EcoRI*. Assim, após a digestão com *EcoRI*, três fragmentos diferentes se hibridizam ao clone de cDNA. Quando o DNA genômico de alguns camundongos albinos é submetido a análise similar, nenhum fragmento de DNA contém sequências complementares ao mesmo cDNA. Esse resultado indica que esses camundongos não têm a habilidade de produzir tirosinase porque o DNA que codifica a enzima deve ter sido deletado.

22. O promotor e as regiões controladoras do gene da planta de interesse devem ser clonados e unidos na orientação correta com o gene de glicuronidase, que coloca o gene repórter sob o mesmo controle transcricional que o gene de interesse. O texto acompanhante descreve a metodologia usada para criar plantas transgênicas. Transformar as células da planta com o gene repórter, e, como discutido no texto, cultivar em plantas transgênicas. O gene de glicuronidase agora será expresso no mesmo padrão desenvolvimental que o do gene de interesse, e sua expressão pode ser facilmente monitorada banhando a planta em uma solução X-Gluc e dosando quanto ao produto azul da reação.

Índice Alfabético

A

Acentuassomo(s), 340
 - β -interferon, 341
 - desoxirribonucleico (ν DNA)
 - nucleicos, 621
 - ribonucleico (ν RNA)
 Adenina, 4
 Adenosina monofosfato cíclica, 312
Aedes aegypti, 411
 Aflatoxina B₁, 452
 Agentes intercalares, 451
Agrobacterium tumefaciens, 633
 Albinismo, 8
 - base molecular do, 9
 Álcool desidrogenase, 590
 Alelo(s), 41, 189
 - detecção de, de doenças humanas, 629
 - do tipo selvagem, 44
 - frequência genotípica para, no locus de grupo sanguíneo MN em várias populações humanas, 519
 - interação gênica de, 192
 - - de um único gene, variações de dominância, 189
 - - - co-dominância, 191
 - - - completa e recessividade, 189
 - - - incompleta, 190
 - - - letais recessivos, 192
 - - no padrão de folhas de trevo, 192
 - múltiplos, 189
 - mutantes, mapeamento de, por pseudodominância, 493
 - natureza dos, e seus produtos, 41
 - nucleares MATa e MAT α , 333
 - nulos, 44
 Alopoliplóides, 486
 - origem de três espécies, de *Brassica*, 482
 Alostéria, evidência genética da, 309
 Alquilação, 451
 Alterações cromossômicas em larga escala, 475-515
 - índice geral de mutações cromossômicas humanas, 503
 - mudanças na estrutura de cromossomos, 490
 - - aplicações de inversões e translocações, 500
 - - deleções, 492
 - - duplicações, 494
 - - identificação das mutações cromossômicas pela genômica, 503
 - - inversões, 495
 - - rearranjos e câncer, 502
 - - translocações, 499
 - - - recíprocas, 499
 - - - robertsonianas, 499
 - mudanças no número de cromossomos, 476
 - - aneuploidia, 484
 - - - monossômicos, 485
 - - - não-disjunção, 484
 - - - trissômicos, 486
 - - conceito de balanço gênico, 488
 - - euploidia aberrante, 476
 - - - animais poliplóides, 483
 - - - aplicações agrícolas, 481
 - - - monoplóides, 477
 - - - poliplóides, 478
 Ames, teste de, 451
 Aminoácidos, 7, 277, 284
 - especificando a sequência de, de uma proteína, 5
 - interações de, 322

- modificação pós-traducional de cadeias laterais de, 293
 - número de substituições de, na evolução dos vertebrados, 589
 - peptídeos-líder de óperons de biossíntese de, 318
 - polimorfismo de sequência de, 520
 - tirosina, 9
 Aminoacil-tRNA sintetase, 283
 2-Aminopurina, 450
 Amniocentese, 630
 Amplificação gênica, como funciona a, 613
 Análise genética, 12, 660
 - *Arabidopsis thaliana*, 654
 - *Caenorhabditis elegans*, 656
 - das tétrades e indícios do mecanismo de recombinação, 464
 - de PEV, 348
 - direta usando clonagem posicional, 626
 - - para identificar um gene de doença humana, 627
 - - para identificar um gene importante para a domesticação do milho, 628
 - do promotor *lac*, 309
 - dos processos bioquímicos da célula, 12
 - *Drosophila melanogaster*, 658
 - *Escherichia coli*, 649
 - *Mus musculus*, 660
 - *Neurospora crassa*, 652
 - *Saccharomyces cerevisiae*, 650
 Anemia falciforme, hemácias de uma pessoa com, 534
 Aneuplóide, 478
 Aneuploidia, 484
 - monossômicos, 485
 - não-disjunção, 484
 - trissômicos, 486
 Anfidiplóide, 480
 Anfidiplóidia ancestral, origem proposta para o trigo do pão por, 483
 Animais, 361
 - domésticos, 553
 - em desenvolvimento, métodos para visualização de expressão gênica nos, 361
 - engenharia genética em, 633
 - poliplóides, 483
Antennapedia, 256
 Anticódon, 6, 283, 285
 Anticorpo, 618
 Antígeno nuclear de proliferação celular, 242
 Apoptose, 467
Arabidopsis thaliana, 14, 26, 293, 654
 - ciclo de vida, 654
 - crescendo na natureza, 654
 - mutantes de, 655
 Arabinose, óperon de, 313
 Arcabouços, 393
 Arginina, 196
 - e seus análogos, 197
 - isolamento de uma mutação auxotrófica para, 196
Astyanu mexicanus, 593
 Ativadores transcricionais eucarióticos, 337
 Atrofia muscular bulbar e espinhal ligada ao X, 448
 Autofecundação, 32
 - dífrida, 92
 - - histograma de poligenes de uma, 93
 - - poligenes na prole de uma, 92
 - trífrida, histograma de poligenes de uma, 93
 Autopoliplóides, 478
 Auto-radiograma, 617

Autotetraplóide, 482
 - pareamento cromossômico em um, 486

B

Bacillus subtilis, esporulação em, 323
 Bactéria(s), 331, 612
 - *Agrobacterium tumefaciens*, 633
 - *Deinococcus radiodurans*, 276
 - em divisão, 151
 - entrada de moléculas recombinantes nas, 615
 - *Haloarcula marismortui*, 275
 - símbolos genotípicos usados em genética de, 154
 - sofrendo transformação, 167
 - *Streptococcus pneumoniae*, 226
 - *Thermus aquaticus*, 622
 - tipo selvagem, 153
 Bactérias, genética de, e seus vírus, 151-186
 - conjugação bacteriana, 154
 - - de bacteriófagos, 167
 - - - infecção por fagos, 168
 - - - mapeamento de cromossomos de fagos usando cruzamentos de fagos, 169
 - - descoberta da conjugação, 155
 - - descoberta do fator fertilidade, 156
 - - linhagens Hfr, 157
 - - mapeamento de cromossomos bacterianos, 161
 - - plasmídeos, 163
 - - - F que levam fragmentos genômicos, 163
 - - - R, 164
 - mapas físicos e mapas de ligação comparados, 174
 - trabalhando com microrganismos, 153
 - transdução, 171
 - - descoberta da, 171
 - - especializada, 173
 - - - mecanismo de, 174
 - - generalizada, 171
 - transformação bacteriana, 167
 Bactérias, regulação da expressão gênica em, e seus vírus, 301-327
 - ciclos de vida de bacteriófagos, 317
 - - anatomia molecular da mudança genética, 320
 - - ligação específica de sequência de proteínas regulatórias ao DNA, 321
 - controle duplo positivo e negativo, o óperon de arabinose, 313
 - descoberta do sistema *lac*, controle negativo, 306
 - - análise genética do promotor *lac*, 309
 - - caracterização molecular do repressor *lac* e operador *lac*, 310
 - - evidência genética, 307
 - - - da alosteria, 309
 - - - do operador e repressor, 307
 - - genes controlados juntos, 307
 - - mutações polares, 311
 - fatores sigma alternativos que regulam grandes conjuntos de genes, 322
 - regulação gênica, 303
 - repressão catabólica do óperon *lac*, controle positivo, 311
 - vias metabólicas e níveis adicionais de regulação, atenuação, 314
 Bacteriófagos, 152
 - ciclo de vida de, 317, 319
 - anatomia molecular da mudança genética, 320
 - ligação específica de sequência de proteínas regulatórias ao DNA, 321
 Balanço gênico, conceito de, 488
 Bandas cromossômicas, 28

Besouros transgênicos, criação de, expressando uma proteína verde fluorescente, 411
 Biblioteca genômica, 392
 - construção de, e de cDNA, 616
 - de clones, 391
 Bioinformática, significado da sequência genômica, 395
 Biologia, o enfoque genético à, 1-24
 - base molecular da informação genética, 4
 - genes, o ambiente e o organismo, 17
 - - determinação ambiental, 17
 - - determinação genética, 17
 - - interação genótipo-ambiente, 18
 - - ruído de desenvolvimento, 19
 - - três níveis de desenvolvimento, 19
 - - uso do genótipo e fenótipo, 18
 - genética e as perguntas da biologia, 2
 - metodologias usadas em genética, 12
 - organismos-modelo, 13
 - programa da investigação genética, 7
 - - começando com o DNA: genética reversa, 11
 - - começando com variação: genética direta, 8
 - - necessidade da variação, 7
 BLAST, pesquisa, 398
 Bloqueios enzimáticos, doenças associadas aos, 198
Bombyx mori, 411
 Braços cromossômicos, 463
Brassica, origem de três espécies aloploplóides de, 482

C

Caçadas genômicas para vários sítios de ligação, 397
 Cadeia polipeptídica, transferência de informação eucariótica do gene para a, 397
Caenorhabditis elegans, 15, 378, 498, 656
 - adulto, 656
 - ciclo de vida, 656
 - e espécies, 379
 - regulação da tradução do mRNA e linhagem celular em, 377
 - transgênese em, 634
 Camundongo(s), 194
 - adulto e sua prole, 660
 - e humanos, genômica comparativa, 400
 - *imprinting* genômico em, 345
 - produção de, contendo o nocaute gênico direcionado, 638
 - sentença entre o cromossomo humano 17 e o cromossomo 11 de, 402
 - terapia gênica em, 639
 - transgênico, produção de um, 661
 Câncer, 502
 - colorretal não-polipose hereditária (*v.* Síndrome HNPCC)
 - como uma doença do desenvolvimento, 382
 - por translocação somática, 503
 - telômeros e envelhecimento, 246
 - uma consequência fenotípica importante da mutação, 467
 Carboxila, domínio da cauda de, 263
 Carga genética, 477
 Castileja, herança quantitativa da bráctea da, 546
Castilleja hispida, 546
 Catabolismo, proteína ativadora do, 312
 Cauda poli(A), 264
 Célula(s), 94, 650
 - análise genética dos processos bioquímicos da, 12
 - bacterianas, maneiras de levar DNA recombinante para, 616
 - cancerosas, 467
 - - diferença das, das células normais, 467
 - - mutações em, 467
 - de *Euglena gracilis*, 94
 - de levedura, 333
 - - micrografia eletrônica de, em brotamento, 333
 - - *Saccharomyces cerevisiae*, 650

- diplóides, 481
 - distorcidas falciformes, 534
 - dobramento da proteína dentro da, 292
 - eucariótica, 6
 - falciforme, 191
 - genoma nuclear nas, de uma fêmea do cervo indiano, 28
 - hospedeira, infecção da, cruzamento de fago feito por dupla, dentro de fagos parentais, 170
 - modos diferentes de introduzir DNA exógeno em uma, 631
 - mostrando nucleóides dentro de mitocôndria, 94
 - normais, 467
 - - e células transformadas por um oncogene, 467
 - - micrografias eletrônicas de, e células transformadas pelo vírus do sarcoma de Rous, 467
 - produtoras contendo um nocaute gênico direcionado, 637
 - um modo para gerar RNA bifilamentar em uma, 270
 Centimorgan, 116
 Centrômero, 28
 - mapeamento de, com tétrades lineares, 127
 Cerdas escutulares de *Drosophila* adulta, 587
Chaetodipus intermedius, 591
 Chargaff, regras de, 229
 Chimpanzés, 11
 - e humanos, genômica comparativa, 402
 - percentagem de nucleotídeos que diferem entre humanos e, 11
 ChIP, método, 407
 Ciclo celular, 35
 - assexual, estágios do, 35
 - de levedura, replicação do DNA e o, 244
 Citosina, 4
 Citotipo, 428
 - M, 428
 - P, 428
Clarkia elegans, 522
 Clonagem, 628
 - de DNA, 610
 - - plasmídeo construído como vetor de, 649
 - de Dolly, 329
 - em fago, 615
 - posicional, 620
 - - análise genética direta usando, 626
 - - - para identificar um gene de doença humana, 627
 - - - para identificar um gene importante para a domesticação do milho, 628
 - uso de elemento P para marcar gene para, 431
 - vetores de, escolha dos, 613
 Clone(s), 394
 - biblioteca genômica de, 391
 - celular, 153
 - *contig*, 394
 - criando um mapa físico por mapeamento de *fingerprint* de, 394
 - específico de interesse, 617
 - sequenciamento genômico de, ordenado, 391
 Cloroplastos, 16
 - mapas de DNA para mitocôndrias e, 94
 Cockayne, síndrome de, 456
 Código(s) genético(s), 234, 279
 - códons de fim, 282
 - decifrando o, 282
 - número de letras no códon, 280
 - redundância do, 282
 - superpostos *versus* não-superpostos, 280
 - uso de supressores para demonstrar um, triplo, 281
 Códon(s), 285
 - de fim, 282
 - de serina, diferentes RNA transportadores que podem servir de, 286
 - de triptofano, 317
 - número de letras no, 280

- tendenciosidade de, 611
 - - previsões baseadas em, 398
 Coeficiente, 571
 - de coincidência, 120
 - de correlação, 571
 - de endogamia, 529
 Colchicina, 479
 Colônias bacterianas, 154, 648
 Complementação, teste de, 199
 - definição de conjuntos de genes usando o, 199
 - usando um heterocáion, 203
 Complexo(s)
 - gênicos, 360
 - MSL, 329
 - SWI-SNF, 338
 Compostos mutagênicos, 453
 Condensação cromossômica por super-helicoidização, 29
 Conjugação bacteriana, 154
 - descoberta da conjugação, 155
 - descoberta do fator fertilidade, 156
 - linhagens Hfr, 157
 - mapeamento de cromossomos bacterianos, 161
 - plasmídeos, 164
 - - F que levam fragmentos genômicos, 163
 - - R, 164
 Constituições cromossômicas em um organismo normalmente diplóide, 477
Contig, sequências, 392
 Conversão genética, 464
 Cosmídeos, 614
 Covariança, 571
 cpDNA (DNA de cloroplastos), 94
Cri du chat, síndrome do, 493
 - origem de deleção da, 494
Crinum longifolium, sementes de, 569
 Cromátides, 113
 - idênticas, 44
 - não-irmãs, 463
 Cromátides-irmãs, 36, 113
 Cromatina, 28, 243
 - dinâmica e regulação gênica, 337
 - domínios de, e sua herança, 345
 - - análise genética de PEV revela proteínas necessárias para a formação de heterocromatina, 348
 - - comparação da heterocromatina e eucromatina, 346
 - - herança de marcas epigenéticas e estrutura cromossômica, 351
 - - mudança do tipo reprodutivo e silenciamento gênico, 346
 - - silenciamento de um cromossomo inteiro, 350
 - - variação do efeito de posição em *Drosophila* revela a vizinhança genômica, 347
 - estrutura da, 338
 - herança de estados da, 352
 - imunoprecipitação de, 406
 - - estudo do interatoma usando, 406
 - - etapas na análise de, 407
 - proteínas remodeladoras de, e ativação gênica, 338
 - remodelagem de, 337
 - - histonas e, 339
 Cromossomo(s), 628
 - andar no, 622
 - - e salto, 628
 - artificial, 614
 - - bacteriano, 614
 - - baseado em P1, 614
 - - de levedura, 614
 - bacterianos, 161
 - - mapeamento de, 161
 - - replicante, 237
 - de milho, 418
 - de tomate, 121
 - distribuição independente de, na meiose, 87
 - genes e, 27

- homólogos, 28
- humano, 400
- - 7, mapa citogenético do, 401
- - 17, sintenia entre o, e o cromossomo 11 de camundongo, 402
- - 20, mapa do, sequência de, 400
- mapeamento de, de fagos usando cruzamentos de fagos, 169
- marcadores fenotípicos e moleculares mapeados no, 128
- mudanças na estrutura de, 490
- - aplicações de inversões e translocações, 500
- - deleções, 492
- - duplicações, 494
- - identificação das mutações cromossômicas pela genômica, 503
- - inversões, 495
- - rearranjos e câncer, 502
- - translocações, 499
- - - recíprocas, 499
- - - robertsonianas, 499
- mudanças no número de, 476
- - aneuploidia, 484
- - - monossômicos, 485
- - - não-disjunção, 484
- - - trissômicos, 486
- - conceito de balanço gênico, 488
- - euploidia aberrante, 476
- - - animais poliplóides, 483
- - - aplicações agrícolas, 481
- - - monoplóides, 477
- - - poliplóides, 478
- politênicos, 492
- problema da replicação nas pontas dos, 245
- sexuais, 51
- silenciamento de um, inteiro, 350
- super-helicoidizado, 29
- translocação recíproca demonstrada por coloração de, 475
- X, inativação do, 350
- Cromossomos eucarióticos, mapeamento de, por recombinação, 109-149
- com marcadores moleculares, 120
- - polimorfismos, 122
- - - de comprimento de sequências simples, 125
- - - de nucleotídeo único, 122
- - cálculo de *crossings* múltiplos não-vistos, 132
- - de centrômero com tétrades lineares, 127
- diagnóstico de ligação, 111
- - como os *crossings* produzem recombinantes para genes ligados, 112
- - evidência de que o *crossing over* é um processo de quebra e reunião, 113
- - evidência de que o *crossing over* ocorre no estágio de quatro cromátides, 113
- - múltiplos *crossings* podem incluir mais de duas cromátides, 114
- - simbolismo de ligação e terminologia, 113
- - uso de frequência de recombinantes para reconhecer ligação, 111
- por frequência, 114
- - cruzamento-teste de três pontos, 117
- - dedução da ordem dos genes por inspeção, 119
- - interferência, 119
- - unidades de mapa, 115
- - uso de proporções como diagnóstico, 120
- uso de mapas baseados em recombinação em conjunto com mapas físicos, 134
- uso de valores Lod para avaliar a ligação em heredogramas humanos, 131
- uso do teste do qui-quadrado para testar análise de ligação, 129
- Crossing(s)*, 113
- como os, produzem recombinantes para genes ligados, 112
- cálculo de, múltiplos não-vistos, 132
- evidência de que o, 113
- - é um processo de quebra e reunião, 113

- - ocorre no estágio de quatro cromátides, 113
- - meiótico, mecanismo de, 463
- - análise genética das tétrades e indícios do mecanismo de recombinação, 464
- - modelo de quebra bifilamentar para recombinação meiótica, 465
- - quebras bifilamentares programadas e a recombinação meiótica, 463
- modelo de decisão precoce de, 466
- múltiplos, podem incluir mais de duas cromátides, 114
- Cruzamento(s), 169
- de fagos, 169
- - feito por dupla infecção da célula hospedeira dentro de fagos parentais, 170
- - mapeamento de cromossomos de fagos usando, 169
- de Mendel, 33
- - programa de, que produziu uma proporção 9:3:3:1, 77
- de *Neurospora*, estágios de um, 88
- diíbridos, 76
- usando flores de uma planta variegada, 96
- Cruzamento-teste, 50
- de três pontos, 117

D

- Datura*, trissômicos de, 488
- Deficiência enzimática, 440
- Deinococcus radiodurans*, bactéria, 276
- Deleção(ões), 492
- alça de, 492
- intragênica, 492
- mapeamento de, 492
- multigênicas, 492
- Deriva genética, 518, 539
- Desenvolvimento, o controle genético do, 355-385
- de *Drosophila*, 357
- - classificação dos genes pela função desenvolvimental, 358
- - genes homeóticos e identidade segmentar, 359
- - grupos de genes *Hox*, 363
- - homeobox, 362
- - organização e expressão dos genes *Hox*, 360
- - definição de todas as ferramentas, 365
- e doença, 381
- enfoque genético, 356
- papéis dos genes ferramentais individuais, 380
- regulação espacial da expressão gênica, 370
- - desenho das faixas, integração dos impulsos de proteínas *gap*, 372
- - gradientes maternos e ativação gênica, 370
- - tornando os segmentos diferentes, integração dos estímulos *hox*, 373
- regulação pós-transcricional da expressão gênica, 374
- - da tradução do mRNA e linhagem celular em *Caenorhabditis elegans*, 377
- - miRNA e *Caenorhabditis elegans* e outras espécies, 379
- - recomposição do RNA e determinação do sexo em *Drosophila*, 376
- - traducional no embrião inicial, 377
- Díade, 36
- Diagnóstico genético molecular, 628
- Dimetil lisina, 349
- Dimorfismo, 58
- herança de um, 60
- Diplóides, 85
- parciais, 307
- síntese de β -galactosidase e permease em linhagens haplóides e, 309
- Disgenesia híbrida, 428
- em *Drosophila*, 429
- Dissecção genética, 26
- Distribuição(ões)
- de Poisson, 132
- fenotípica(s), 551
- - de genótipos de três plantas, 550
- - genótipos e, 549
- - norma de reação e, 551
- Distúrbio(s)
- autossômico(s), 55
- - dominante, 56
- - - herança de um, 57
- - ligados ao X, 59, 61
- - - herança de um, 62
- - recessivos, 55
- genéticos, 401
- - heterogêneos, 458
- - pontos de quebra de rearranjos de pacientes com, 401
- DNA, 225-251
- amplificação de, *in vitro*, a reação em cadeia da polimerase, 622
- análise direta do, 12
- bacteriano, 153
- - transdução generalizada por incorporação aleatória de, em cabeças de fagos, 172
- clonagem de, 610
- - plasmídeo construído como vetor de, 649
- complementar, 611
- - de um tecido que expressa um gene de interesse em alto nível, 617
- cromossômico enrolado ao redor de histonas, 29
- danificado, reversão direta do, 454
- de cloroplastos, 94
- de cópia única, 392
- determinando a sequência de bases de um segmento de, 624
- divisão nuclear ao nível de, 45
- doador, 610
- - ligação do, e vetor, 612
- - tipo de, 611
- - transferido como um filamento único, 158
- domínio de ligação ao, 304
- dupla hélice de, 5
- - duas representações da, 233
- duplicação do, 595
- e transmissão gênica na mitose e meiose em eucariontes, 45
- estrutura do, 228
- - antes de Watson e Crick, 229
- - de dupla hélice, 231
- - e polimorfismo de sequências, 521
- estrutura dos agentes intercalares comuns e sua interação com o, 451
- estruturas dos quatro nucleotídeos do, 230
- estruturas dos sítios-alvo de, 312
- exógeno, modos diferentes de introduzir, em uma célula, 631
- filamentos opostos de, 257
- *fingerprints* de, 125
- genômico, cortando o, 611
- girase, 241
- herança monogênica rastreada ao nível de, 46
- importado, 597
- ligação específica de sequência de proteínas regulatórias ao, 321
- ligase, 238
- - enzima, 612
- mapas de, para mitocôndrias e cloroplastos, 94
- material genético, 226
- - descoberta da transformação, 226
- - experimento de Hershey-Chase, 227
- metilação do, 344
- microarranjos de, 405
- - padrões de expressão gênica detectados por, 406
- - uso de, para estudar o transcriptoma, 405
- mitocondrial, 94
- - sítios de mutações no, em algumas doenças humanas, 98
- modelo computadorizado do, 225
- modelo de herança da metilação do, 352
- moléculas de, 44

- - específicas, detecção de, 12
- - modelo de Watson-Crick da, 226
- mudanças estruturais no, 497
- mudanças tautoméricas nas bases do, 446
- natureza do conteúdo de informação do, 395
- nuclear universal, código de, comparação do, com vários códigos mitocondriais, 598
- palíndromo de, 611
- pareamento de bases no, 233
- polimerase, 237
- - reação catalisada pela, 237
- - *taq* polimerase, 622
- primeiro modelo de, 231
- promotor, 303
- recombinação de cassetes de, 347
- recombinante, 610
- - inserção de um gene em um plasmídeo de, 612
- - maneiras de levar, para células bacterianas, 616
- - moléculas de, 610
- - - formação de uma, 612
- - técnicas de, 620
- replicação de, 5
- - dispersiva, 235
- - e o ciclo celular de levedura, 244
- - em organismos eucarióticos, 241
- - erro na, 445
- - montagem de nucleossomos durante a, 243
- - origens eucarióticas, 242
- - - em eucariontes superiores, 244
- - replissomo, 239
- - - deselicioidizando a dupla hélice, 240
- - - eucariótico, 242
- - - montagem do, 241
- - semiconservativa, 234
- - - experimento de Meselson-Stahl, 235
- - - forquilha de replicação, 235
- - - polimerases, 236
- - telômeros e telomerase, 245
- - visão geral da, 238
- sequência de, 396
- - curta, 424
- - e RNA transcrito, 259
- - evidência direta, 396
- - leitura da, de um sequenciador automático, 626
- sinal característico da seleção positiva nas sequências de, 590
- sinal característico da seleção purificadora no, 588
- sintetizado quimicamente, 611
- sítios de ligação de, 397
- sondas de, 618
- - específico, 12
- tecnologia do, 152, 610
- translesão, síntese de, 459
- transposons de, utilidade dos, para a descoberta do gene, 428
- uso de polipeptídeo e similaridade de, 396
- Doença(s)
- associadas aos bloqueios enzimáticos, 198
- de Huntington, 448
- - de manifestação tardia, 58
- de imunodeficiência combinada grave (*v.* Doença do menino da bolha)
- de Kennedy, 448
- de repetições de trinucleotídeos, 448
- desenvolvimento e, 381
- detecção de alelos de, e diagnóstico genético molecular, 628
- do menino da bolha, 417
- genéticas comuns, 630
- herdada recessivamente, 440
- identificando um gene de, 627
- mitocondrial, heredograma de uma, 98
- sítios de mutações no mtDNA em algumas, 98
- Dolly, 330
- clonagem de, 329
- Dosagem gênica, efeito da, 489
- Down, síndrome de, 476, 486

- características da, 487
- efeito da idade materna na, 487
- na prole de uma translocação heterozigota, 501
- Droga eritromicina, 276
- Drosophila*, 347, 358
- alças de deleção em, 492
- *biarmipes*, machos de, 593
- cerdas escutelares de, adulta, 587
- desenvolvimento de, 359
- disgenesia híbrida em, 429
- dois mutantes morfológicos de, com o tipo selvagem para comparação, 658
- embriões de, 368
- ferramentas para o desenvolvimento de, 357
- - classificação dos genes pela função desenvolvimental, 358
- - genes homeóticos e identidade segmentar, 359
- - organização e expressão dos genes *Hox*, 360
- heterozigota, linhagem de, 493
- *melanogaster*, 14, 52, 357, 658
- - ciclo de vida, 658
- - machos de, 593
- - mutantes homeóticos de, 356
- - transgênese em, 634
- mutações que interagem e afetam a cor do olho em, 10
- polimorfismos sinônimos e não-sinônimos e diferenças de espécies, 590
- proteína *Hox* de, e de vertebrados, 363
- *pseudoobscura*, 521
- - viabilidades de vários cromossomos homozigotos de, em três temperaturas diferentes, 534
- recomposição do RNA e determinação do sexo em, 376
- segmentos torácicos e abdominais normais de, 659
- variedade de alças de inversão encontradas em populações de, 522
- *willistoni*, 569
- Duplicações cromossômicas, 494

E

- Efeito do fundador, 539
- Efetores alostéricos, 304
- Elemento P, uso de, 428
- para inserir genes, 431
- para marcar gene para clonagem, 431
- Eletroforese em gel, 619
- Embrião(ões)
- animais, organizadores em, 357
- controle genético traducional no, inicial, 377
- da mosca-da-fruta, expressão gênica de um desenvolvimento de, 355
- de *Drosophila*, 368
- Embrióide, 481
- Endogamia, 518, 528, 551, 564
- coeficiente de, 529
- e reprodução preferencial, 528
- heterozigose após várias gerações de, para dois sistemas de reprodução, 554
- homozigotos recessivos de, 56
- negativa, 528
- Enfoque *shotgun*, sequenciamento de um genoma usando o, de genoma inteiro, 392
- para criar uma sequência-rascunho de um genoma complexo, 392
- simples, 392
- Engenharia genética, 610, 631
- animais, 633
- *Arabidopsis thaliana*, 655
- *Caenorhabditis elegans*, 657
- *Drosophila melanogaster*, 659
- *Escherichia coli*, 649
- humanos, 639
- *Mus musculus*, 661

- *Neurospora crassa*, 653
- plantas, 631
- *Saccharomyces cerevisiae*, 631, 651
- Envelhecimento, 246
- prematuro, 247
- telômeros, câncer e, 246
- Enzima(s), 9, 611
- β -galactosidase, 305
- de restrição, 12, 611
- distributiva, 240
- DNA ligase, 612
- processiva, 240
- que funciona mal e albinismo, 9
- SpoII, 463
- tirosinase, 9
- Epistasia, 204
- de sinal, 586
- dominante, 206
- recessiva, 204
- - devida a mutação de pelagem amarela, 205
- Equilíbrio
- de Hardy-Weinberg, 525
- genético, segregação meiótica e, 525
- Eritromicina, 276
- Erros
- hereditários do metabolismo, descobridor dos, 195
- na replicação do DNA, 445
- Escherichia coli*, 14, 155, 259, 586, 648
- ciclo de vida, 648
- genoma de, 648
- - mapa do, 176
- genômica comparativa de, não-patogênica e patogênica, 403
- linhagens de, 404
- reparo de mal pareamento em, 459
- Especiação alopátrica, processo de, 603
- Espécies dióicas, 51
- Esporos, descobrindo um gene para a produção de, 49
- Estudo(s)
- de interações de proteínas, com o uso do sistema díbrido de levedura, 407
- do interatoma, 406
- do transcriptoma, 405
- Etilmetanossulfato, 450
- Eucariontes, 14, 289
- DNA e transmissão gênica na mitose e meiose em, 45
- elementos de transposição em, 425
- início da tradução em, 289
- multicelulares, 342
- regulação da expressão gênica em, 329-354
- - cromatina dinâmica e, 337
- - domínios de cromatina e sua herança, 345
- - - análise genética de PEV e proteínas necessárias para a formação de heterocromatina, 348
- - - comparação da heterocromatina e eucromatina, 346
- - - herança de marcas epigenéticas e estrutura cromossômica, 351
- - - mudança do tipo reprodutivo e silenciamento gênico, 346
- - - silenciamento de um cromossomo inteiro, 350
- - - variação do efeito de posição em *Drosophila* revela a vizinhança genômica, 347
- - *imprinting* genômico, 343
- - lições de leveduras, o sistema GAL, 334
- - mecanismo de ação acentuada, 340
- - transcricional, 330
- superiores, origens de replicação do DNA em, 244
- transcrição do RNA em, 261
- - alongamento, término e processamento pré-mRNA, 263
- - início, 262

Eucromatina, 28, 350
 - comparação da heterocromatina e, 346
Euglena gracilis, célula de, 94
 Euplóides, 476
 Euploidia aberrante, 476
 - animais poliplóides, 483
 - aplicações agrícolas, 481
 - monoplóides, 477
 - poliplóides, 478
 Evidência genética, 307
 - da alosteria, 309
 - de ancestralidade comum na evolução, 598
 - - comparação dos genomas entre vizinhos próximos, 601
 - - comparando os proteomas entre espécies distantes, 600
 - do operador e repressor, 307
 Exogamia forçada, 528
 Éxons, 254
 - mecanismo de recomposição dos, 266
 - reações na recomposição de, 267
 Experimento(s), 418
 - de Griffith, 226
 - de Hershey-Chase, 227
 - de McClintock, 418
 - de Meselson-Stahl, 235
 - de pulso-caça, 254
 - de Rosalind Franklin, 231
 Expressão gênica, 370
 - métodos para visualização de, nos animais em desenvolvimento, 361
 - padrões de, detectados por microarranjos de DNA, 406
 - regulação espacial da, no desenvolvimento, 370
 - - desenho das faixas, integração dos impulsos de proteínas *gap*, 372
 - - gradientes maternos e ativação gênica, 370
 - - tornando os segmentos diferentes, integração dos estímulos *hox*, 373
 - regulação pós-transcricional da, 374
 - - da tradução do mRNA e linhagem celular em *Caenorhabditis elegans*, 377
 - - miRNA e *Caenorhabditis elegans* e outras espécies, 379
 - - recomposição do RNA e determinação do sexo em *Drosophila*, 376
 - - traducional no embrião inicial, 377
 Expressão gênica, regulação da, em bactérias e seus vírus, 301-327
 - ciclos de vida de bacteriófagos, 317
 - - anatomia molecular da mudança genética, 320
 - - ligação específica de sequência de proteínas regulatórias ao DNA, 321
 - controle duplo positivo e negativo, o óperon de arabinose, 313
 - descoberta do sistema *lac*, controle negativo, 306
 - - análise genética do promotor *lac*, 309
 - - caracterização molecular do repressor *lac* e operador *lac*, 310
 - - evidência genética, 3-7
 - - - da alosteria, 309
 - - - do operador e repressor, 307
 - - genes controlados juntos, 307
 - - mutações polares, 311
 - fatores sigma alternativos que regulam grandes conjuntos de genes, 322
 - regulação gênica, 303
 - repressão catabólica do óperon *lac*, controle positivo, 311
 - vias metabólicas e níveis adicionais de regulação, atenuação, 314
 Expressão gênica, regulação da, em eucariontes, 329-354
 - cromatina dinâmica e, 337
 - domínios de cromatina e sua herança, 345
 - - a análise genética de PEV revela proteínas necessárias para a formação de heterocromatina, 348

- - comparação da heterocromatina e eucromatina, 346
 - - de marcas epigenéticas e estrutura cromossômica, 351
 - - mudança do tipo reprodutivo e silenciamento gênico, 346
 - - silenciamento de um cromossomo inteiro, 350
 - - variação do efeito de posição em *Drosophila* revela a vizinhança genômica, 347
 - *imprinting* genômico, 343
 - lições de leveduras, o sistema GAL, 334
 - mecanismo de ação acentuada, 340
 - transcricional, 330

F

Fago(s), 152
 - ciclo de um, 169
 - clonagem em, 615
 - cruzamento de, 169
 - - feito por dupla infecção da célula hospedeira dentro de fagos parentais, 170
 - - mapeamento de cromossomos de fago usando, 169
 - infecção de bactérias por, 168
 - material genético do, 229
 - recombinantes, placas de prole de, e parentais, 170
 - transdução generalizada por incorporação aleatória de DNA bacteriano em cabeças de, 172
 Familiaridade e herdabilidade, 556
 Feminização testicular, fenótipo de, 62
 Fenilalanina, via metabólica de, 198
 Fenocópia, genética reversa por, 409
 Fenótipo(s), 31, 553
 - albino, 8
 - de feminização testicular, 62
 - de pseudo-acondroplasia humana, 57
 - e diferença em ambientes diversos, 553
 - estudos dos, por Mendel, 31
 - incomuns, 420
 - instável, 421
 - liso e rugoso, 77
 - poky, mutante mitocondrial de, 95
 - produzidos por elementos de transposição em grãos de milho, 421
 - uso do genótipo e, 18
 Filamento(s)
 - *lagging*, síntese do, 239
 - helicoidização de, dependente de síntese, 462
Fingerprint, 125
 - de DNA, 125
 - mapeamento de, criando um mapa físico, de clones, 394
 Flores, 49
 - compostas de *Gaillardia pulchella*, 545
 - conjunto de genes instrumentais no desenvolvimento de, 49
 - cruzamentos usando, de uma planta variegada, 96
 - da planta campainha, 200
 - de petúnia demonstrando co-supressão, 269
 Flutuação, teste de, de Luria-Delbrück, 443
 Folhas variegadas causadas por uma mutação em cpDNA, 96
 Fórmula de Perkins, 133
 Fosfatase ácida, atividade das hemácias de genótipos diferentes de, 566
 Fotoprodutos gerados por luz ultravioleta, 452
 Fragmento(s)
 - de Okazaki, 238, 245
 - de restrição, 611
 - genômicos, plasmídeos F que levam, 163
 Frequência(s)
 - gênica, taxa de mudança em, 535
 - genotípicas, 518

- - para alelos no locus de grupo sanguíneo MN em várias populações humanas, 519
 Função
 - gênica, perturbação da, com o uso de RNA de interferência, 409
 - protéica, perturbação da, com o uso de genética química, 410
 Fungo(s), 14
 - *Neurospora crassa*, 89
 - filamentos, 14

G

Gaillardia pulchella, flores compostas de, 545
 Galactose, 334
 Gametas, formação de, 78
 Gel, eletroforese em, 619
 Gene(s)
 - *bcd*, 366
 - candidato, 565
 - cI, 318
 - cII, 318
 - cIII, 318
 - classificação dos, pela função desenvolvimental, 358
 - codificantes, deduzindo os, de proteína da sequência genômica, 395
 - colinearidade de, e proteína, 279
 - conjunto de, instrumentais, 49
 - - no desenvolvimento de flores, 49
 - - no desenvolvimento de hifas, 50
 - consequências de mutações de ponto dentro de, 441
 - controlados coordenadamente, 307
 - controlados juntos, 307
 - *cro*, 318
 - de efeito materno, 365
 - de manutenção, 357
 - de polaridade segmentar, 366
 - de regra dos pares, 366
 - de resistência, 424
 - definição de conjuntos de, usando o teste de complementação, 199
 - descoberta dos, pela observação das proporções de segregação, 47
 - - ativo no desenvolvimento da cor da flor, 48
 - - genética direta, 50
 - - para a produção de esporos, 49
 - - para desenvolvimento das asas, 48
 - - prever as proporções da prole ou genótipos, aplicando os princípios da herança monogênica, 50
 - - resultados da descoberta, 49
 - *DinB*, 460
 - *distal-less*, 374
 - distribuição independente de, 75-107
 - - base cromossômica da, 85
 - - - autossômicos e ligados ao X, 86
 - - - em organismos diplóides, 85
 - - - em organismos haplóides, 86
 - - - recombinação, 89
 - - de organelas, herança independente do núcleo, 93
 - - - mutações citoplasmáticas em humanos, 97
 - - - padrões de, 95
 - - - segregação citoplasmática, 95
 - - herança poligênica, 91
 - - lei de Mendel, 76
 - - trabalhando com a, 79
 - - - prevendo as proporções na prole, 80
 - - - síntese de linhagens puras, 83
 - - - uso do teste de qui-quadrado nas proporções monoíbridas e díbridas, 81
 - - - vigor híbrido, 84
 - *doublesex*, 376
 - *Duffy*, 595
 - e características quantitativas, 546

- e cromossomos, 27
- endógeno, 269
- estruturais *lac*, 304
- eucariótico, promotor de um, 332
- *even-skipped*, 372
- fatores sigma alternativos que regulam grandes conjuntos de, 322
- ferramentais, 366
- expressão dos, 366
- individuais, papéis dos, 380
- *FMR-1*, 448
- GAL, 334
- *gap*, 366, 370
- haploinsuficientes, 46
- haplossuficientes, mutações de, 189
- *hedgehog*, 380
- *HO*, 346
- homeóticos e identidade segmentar, 359
- homólogo(s), 400
- de um organismo correlato, 617
- *Hox*, 373
- grupos de, 363
- organização e expressão dos, 360
- humanos, distribuição de, em outras espécies, 601
- *hunchback*, 371
- identificando um, 627
- de doença humana, 627
- importante para a domesticação do milho, 628
- imprintados, herança incomum de, 345
- inserção de um, em um plasmídeo de DNA recombinante, 612
- interação de alelos de um único, variações de dominância, 189
- alelos letais recessivos, 192
- co-dominância, 191
- completa e recessividade, 189
- incompleta, 190
- *ISL1*, 403
- isolamento e manipulação de, 609-645
- amplificação de DNA *in vitro*, a reação em cadeia da polimerase, 622
- análise genética direta usando clonagem posicional, 626
- para identificar um gene de doença humana, 627
- para identificar um gene importante para a domesticação do milho, 628
- detecção de alelos de doenças humanas, diagnóstico genético molecular, 629
- determinando a sequência de bases de um segmento de DNA, 624
- engenharia genética, 631
- em animais, 633
- em plantas, 631
- em *Saccharomyces cerevisiae*, 631
- humana, 639
- gerando moléculas de DNA recombinante, 610
- amplificação dentro de uma bactéria, 612
- construção de bibliotecas genômicas e de cDNA, 616
- cortando o DNA genômico, 611
- encontrar um clone específico de interesse, 617
- entrada de moléculas recombinantes nas bactérias, 615
- ligação do DNA doador e vetor, 612
- recuperação de moléculas recombinantes amplificadas, 615
- tipo de DNA doador, 611
- *let-7*, 380
- ligados e herdados juntos, 111
- *lin-4*, 380
- localizando os, 565
- análise de ligação quantitativa, 567
- protocolo experimental para localizar, 567
- segregação de genes marcadores, 566
- marcadores, 566
- mutantes, 8
- de segmentação, 367
- e albinismo, 8
- o ambiente e o organismo, 17
- determinação ambiental, 17
- determinação genética, 17
- interação genótipo-ambiente, 18
- ruído de desenvolvimento, 19
- três níveis de desenvolvimento, 19
- uso do genótipo e fenótipo, 18
- ordem do, no óperon *trp*, 316
- ortólogos, 400
- PAH, 189
- parálogos, 400
- PKU, sítios mutantes no, 46
- plasmídeo R com, de resistência levados em um transposon, 166
- protocolo experimental para localizar, 567
- pseudoligação de, em uma translocação heterozigota, 501
- repórter, 335
- sequência de, e proteínas, 280
- *Shh*, 381
- *Sonic hedgehog*, 382
- SRY, 62
- supressores tumorais, 468
- transferência de informação eucariótica do, para a cadeia polipeptídica, 397
- *trp*, 314
- *ultrabithorax*, 256
- *UmuC*, 460
- uso de elemento P para marcar, para clonagem, 431
- utilidade dos transposons de DNA para a descoberta do, 428
- *white*, 429
- *yellow*, 594
- Genética, 2, 12, 151-186, 517-607
- de bactérias e seus vírus, 151-186
- conjugação bacteriana, 154
- descoberta da conjugação, 155
- descoberta do fator fertilidade, 156
- linhagens Hfr, 157
- mapeamento de cromossomos bacterianos, 161
- plasmídeos F que levam fragmentos genômicos, 163
- plasmídeos R, 164
- de bacteriófagos, 167
- infecção de bactérias por fagos, 168
- mapeamento de cromossomos de fagos usando cruzamentos de fagos, 169
- mapas físicos e mapas de ligação comparados, 174
- símbolos genotípicos, 154
- trabalhando com microrganismos, 153
- transdução, 171
- descoberta da, 171
- especializada, 173
- generalizada, 171
- transformação bacteriana, 167
- de populações, 517-544
- efeitos da reprodução sexual na variação, 525
- endogamia e reprodução preferencial, 528
- heterozigose, 527
- reprodução aleatória, 527
- segregação meiótica e equilíbrio genético, 525
- eventos aleatórios, 538
- fontes de variação, 530
- por migração, 532
- por mutação, 530
- por recombinação, 531
- polimorfismo balanceado, 537
- seleção, 533
- como funciona, 534
- duas formas de, 533
- medindo as diferenças de adaptabilidade, 534
- taxa de mudança em frequência gênica, 535
- variação e sua modulação, 518
- estrutura do DNA e polimorfismo de seqüências, 521
- observações da variação, 518
- polimorfismos de proteínas, 520
- direta, 8, 50
- do tipo reprodutivo, 651
- e as perguntas da biologia, 2
- e genômica, recursos de bioinformática para, 665
- evolutiva, 579-607
- darwiniana, 580
- evidência genética de ancestralidade comum na evolução, 598
- comparando os genomas entre vizinhos próximos, 601
- comparando os proteomas entre espécies distantes, 600
- evolução filética, 581
- mutação e evolução molecular, 588
- origem de novos genes, 595
- processo de especiação, 601
- regulatória, 593
- em humanos, 595
- relação entre genética e mudança funcional na evolução de proteínas, 590
- inativação gênica, 592
- morfológica, 590
- sinal característico da seleção positiva nas seqüências de DNA, 590
- uma síntese de forças, variação e divergência de populações, 581
- variação genética, 586
- dentro de populações e entre populações, 588
- herdabilidade, 586
- vários picos adaptativos, 584
- metodologias usadas em, 12
- mitocondrial, 651
- quantitativa, 545-577
- algumas noções estatísticas básicas, 548
- determinando as normas de reação, 552
- estudos de populações naturais, 554
- plantas e animais domésticos, 553
- resultados dos estudos, 554
- genes e características quantitativas, 546
- genótipos e distribuição fenotípica, 549
- herdabilidade, quantificando a, 556
- estimando os componentes da variância genética, 562
- métodos para estimar H^2 , 558
- restrita, 561
- seleção artificial, 563
- significado de H^2 , 559
- uso de H^2 em endogamia, 564
- localizando os genes, 565
- análise de ligação quantitativa, 567
- segregação de genes marcadores, 566
- norma de reação e distribuição fenotípica, 551
- química, perturbação da função proteica com o uso de, 410
- reversa, 408
- e mutagênese, 408
- aleatória, 408
- direcionada, 409
- por fenocópia, 409
- Genoma(s), 30, 601
- comparação dos, entre vizinhos próximos, 601
- de *Escherichia coli*, 648
- mapa físico do, 176
- de mamíferos, 401
- de organela, 94
- de quatro espécies diferentes, 30
- dinâmico, elementos de transposição, 417-438
- abrigos seguros, 434
- descoberta dos, em milho, 418
- autônomos e não-autônomos, 421
- experimentos de McClintock, 418
- do genoma humano, 433
- em eucariontes, 425

- - em procariontes, 422
- - genomas grandes, 432
- - gramíneas, 434
- - mais do que se imaginava, 431
- diplóide, 28
- duplicações no, de *Saccharomyces* por poliploidia ancestral, 495
- elementos de transposição de, 432
- - grandes, 432
- - humano, 433
- eucarióticos, comparação estrutural de um retrovírus com retrotransposons encontrados em, 427
- humano, 399
- - estrutura do, 399
- - mapas, 495
- - - de duplicações segmentares, 495
- - - de sintenia, 601
- nuclear, 27
- - nas células de uma fêmea do cervo indiano, 28
- seqüenciamento genômico *shotgun* de, inteiro, 391
- - estratégias de montagem de, 393
- Genoma e genômica, 387-415
- bioinformática, significado da seqüência genômica, 395
- comparativa, 400
- - de camundongos e humanos, 400
- - de chimpanzés e humanos, 402
- - de *Escherichia coli* não-patogênica e patogênica, 403
- - elementos conservados e ultraconservados não-codificantes, 402
- criando o mapa de seqüência, 389
- - estabelecendo uma biblioteca genômica de clones, 391
- - lógica de criar um mapa, 390
- - preenchendo espaços na seqüência, 395
- - transformando leituras de seqüências em um mapa de seqüência, 389
- - usando o enfoque de clone ordenado para seqüenciar um genoma complexo, 394
- - usando o enfoque *shotgun* de genoma inteiro, 392
- - - para criar uma seqüência-rascunho de um genoma complexo, 392
- funcional e genética reversa, 405
- revolução genômica, 388
- Genômica, 610 (v. tb. Genoma e genômica)
- comparativa humano-camundongo, 601
- funcional com organismos não-modelo, 410
- genética e, recursos de bioinformática para, 665
- Genótipo(s), 156
- atividade das hemácias de, diferentes de fosfatase ácida na população inglesa, 566
- bacterianos, mistura de, 156
- distribuições fenotípicas de, 549
- - de três plantas, 550
- uso do, e fenótipo, 18
- Glicose-1-fosfato, 334
- Glicose, níveis de, e controle do óperon *lac*, 312
- Globinas, 596
- Gradientes maternos e ativação gênica, 370
- Gramíneas, 434
- Grãos, ambiente e produtividade em, 555
- Griffith, experimentos de, 226
- Grupos sanguíneos, 519
- ABO, 520
- alélicos, diferenciação extrema em frequências de, em três grandes grupos geográficos, 587
- MN, 519
- Guanina, 4, 451

H

Haemophilus influenzae, 388
Haloarcula marismortui, bactéria, 275

Haplóides, 85

- herança monogênica em, 37
- síntese de β -galactosidase e permease em, e mutantes operadores diplóides heterozigotos, 308

Haplótipos, uso de, para deduzir posição gênica, 125

Hardy-Weinberg, equilíbrio de, 525

Helicases, 241

Helicoidização de filamento dependente de síntese, 462

Hemácias, 534

- atividade das, de genótipos diferentes de fosfatase ácida na população inglesa, 566
- falciformes, 534

Hemizigotos, 51

Hemofilia, herança de, na realeza européia, 61

Hemoglobina, 191, 278, 596, 597

- A, 17
- S, 17
- separação eletroforética revelando variantes de, 521

Herança, 60, 351

- de estados da cromatina, 352
- de hemofilia, 61
- de um dimorfismo, 60
- dominante, 57
- epigenética, 344, 351
- incomum de genes imprintados, 345
- ligada ao X, 60, 62
- ligada ao Y, 62
- materna de mutante mitocondrial de fenótipo poky, 95
- monoalélica, 344
- poligênica, 91
- quantitativa da bráctea da castileja, 546

Herança monogênica, 25-74

- análise de heredogramas, 53
- - cálculo de riscos na, 63
- - distúrbios, 55
- - - dominantes, 56
- - - ligada ao X, 59, 61
- - - ligada ao Y, 62
- - - recessivos, 55
- - polimorfismos autossômicos, 58
- descobrindo genes pela observação das proporções de segregação, 47
- - ativo no desenvolvimento da cor da flor, 48
- - genética direta, 50
- - para a produção de esporos, 49
- - para desenvolvimento das asas, 48
- - prever as proporções da prole ou genótipos parentais, 50
- - resultados da descoberta, 49
- genes e cromossomos, 27
- padrões de, 30
- - base cromossômica dos, 35
- - - em haplóides, 37
- - - molecular da segregação e expressão monogênica, 38
- - ligada ao sexo, 50
- rastreada ao nível de DNA, 46

Herdabilidade, 556

- da variação genética, 586
- de uma característica quantitativa, 556
- do QI, 557
- familiaridade e, 556
- método padrão para testar, em organismos experimentais, 558
- quantificando a, 557
- - estimando os componentes da variância genética, 562
- - métodos para estimar H^2 , 558
- - restrita, 561
- - seleção artificial, 563
- - significado de H^2 , 559
- - uso de H^2 em endogamia, 564

Heredograma(s), 61

- análise de, 53
- - cálculo de riscos na, 63
- - distúrbios, 55
- - - dominantes, 56, 61
- - - recessivos, 55, 59
- - herança ligada ao Y, 62
- - polimorfismos autossômicos, 58
- de uma doença mitocondrial, 98
- humanos, uso de valores Lod para avaliar a ligação em, 131

Hershey-Chase, experimento de, 227

Heterocáron, 201

- teste de complementação usando um, 203

Heterocromatina, 28, 502

- comparação da, e eucromatina, 346
- constitutiva, 347
- formação de, 348
- proteína-1 de, 349
- silenciamento genético causado pela expansão da, 348

Heterozigose, 527

- após várias gerações de endogamia para dois sistemas de reprodução, 554

Heterozigoto(s), 34, 192

- híbridos, 85

Hibridização, 611

- *in situ*, 361

Hifas, conjunto de genes instrumentais no desenvolvimento de, 50

Histona(s), 28

- acetilação e desacetilação de, e expressão gênica, 339
- acetiltransferase, 350
- cauda de, 339
- - modificadas, 339
- código de, 339
- desacetilase, 340
- DNA cromossômico enrolado ao redor de, 29
- e remodelagem de cromatina, 339
- hiperacetiladas, 340
- octâmero de, 338

HNPCC, síndrome, 459

Holliday, junção de, 439, 465

Holoenzima polimerase III, 239

Holoprosencefalia, 382

Homeobox, 362

Homo sapiens, 30, 254

Homologia, sistemas de reparo genômico dependente de, 454

Homozigotos recessivos, 34

- de endogamia, 56

Hospedeiros bacterianos, plasmídeos com segmentos de muitos, 166

Huntington, doença de, 448

- de manifestação tardia, 58

Hyalophora cecropia, 39

I

Imprinting genômico, 343

- em camundongo, 345

Imunodeficiência, doença de, combinada grave (v. Doença do menino da bolha)

Imunoprecipitação de cromatina, 406

- estudo do interatoma usando, 406
- etapas na análise de, 407

Inativação gênica, 592

Indol-3-glicerol fosfato, 279

Infecção(ões)

- de bactérias por fagos, 168
- de célula hospedeira, cruzamento de fagos feito por dupla, dentro de fagos parentais, 170
- pelo plasmídeo Ti, 633

Informação genética, base molecular da, 4

Insetos não-modelo, 411

- exemplos de, que expressam um transgene, 411
- proteína repórter verde fluorescente transgênica, expressão nos olhos de alguns, 411

Interação

- gênica, 187-224
- de alelos de um único gene, variações de dominância, 189
- - - alelos letais recessivos, 192
- - - co-dominância, 191
- - - completa e recessividade, 189
- - - incompleta, 190
- - em vias, 195
- - - biossintéticas em *Neurospora*, 195
- - - tipos de vias, 197
- - interferentes, 198
- - - análise de mutantes duplos por mutações aleatórias, 201
- - - definição de conjuntos de genes usando o teste de complementação, 199
- - penetrância e expressividade, 209
- genótipo-ambiente, 18

Interatoma, 294

- estudo de, 405

- teste diíbrido, 405

- usando triagem de imunoprecipitação de cromatina, 406

Interruptores genéticos, 303

Íntrons, 266

- auto-recomposição de, 267

- de auto-remoção, 267

Inversões cromossômicas, 495

Investigação genética, programa da, 7

- começando, 8

- - com o DNA: genética reversa, 11

- - com variação: genética direta, 8

- necessidade da variação, 7

Isoformas protéicas, 292

Isolamento, 196, 609

- de uma mutação auxotrófica para arginina, 196

- e manipulação de gene, 609-645

- - amplificação de DNA *in vitro*, a reação em cadeia da polimerase, 622

- - análise genética direta usando clonagem posicional, 626

- - detecção de alelos de doenças humanas, diagnóstico genético molecular, 629

- - determinando a sequência de bases de um segmento de DNA, 624

- - engenharia genética, 631

- - gerando moléculas de DNA recombinante, 610

J

Junção, 439

- de Holliday, 439, 465

- de pontas não-homólogas, 461

K

Kennedy, doença de, 448

Klinefelter, síndrome de, 486

- características da, 487

L

Lactose, 306, 310

- metabolismo da, 305

Lava, pedras de, 591

Lei de Mendel, 30, 76

Leitura de sequenciamento, 390

Levedura(s), 14, 333

- células de, 333

- - micrografia eletrônica de, em brotamento, 333

- - *Saccharomyces cerevisiae*, 650

- ciclo celular de, replicação do DNA e o, 244

- controle do tipo reprodutivo de, 342

- cromossomo artificial de, 614

- do pão, ciclo de vida, 333

- plasmídeo integrativo de, 651

- sistema diíbrido de, estudo de interações de

- proteína com o uso de, 407

- vetores de, simples, 632

Ligação peptídica, 277

Liguus fasciatus, polimorfismo de cor de conchas

em, 517

Linhagem(ns)

- celular em *Caenorhabditis elegans*, regulação da

tradução do mRNA e, 377

- de *Drosophila* heterozigota, 493- de *Escherichia coli*, 404- de *Nicotiana longiflora*, 547

- germinativa, terapia gênica de, 640

- haplóides, síntese de β -galactosidase e

permease em, e diplóides, 309

- Hfr, 157

- homozigotas, 530

- puras, 32

- - síntese de, 83

Lisina, 349

Loci, 110

- de características quantitativas, 91, 566

Luria-Delbrück, teste de flutuação de, 443

Luz ultravioleta, 451

- fotoprodutos gerados por, 452

M

Mamíferos, 401

- genoma de, 401

- terapia gênica em, 640

Manchas, 59

- de piebaldismo dominante, 59

- nas asas da mosca-das-frutas, 593

Manipulação genética, técnicas de, 651

- *Neurospora crassa*, 653- *Saccharomyces cerevisiae*, 651

Mapa(s)

- cromossômico, 110, 160

- de sequência genômica (*v.* Sequência genômica,

mapa de)

- de sintenia, 601

- - camundongo-humano dos 12 cromossomos do

genoma humano, 660

- - do genoma humano, 601

- do cromossomo, 400

- - humano 20, sequência de, 400

- - humano 7, citogenético, 401

Mapeamento, 394

- de alelos mutantes por pseudodominância, 493

- de cromossomos, 161

- - bacterianos, 161

- - de fagos usando cruzamentos de fagos, 169

- de *fingerprint*, criando um mapa físico por, de

clones, 394

Mapeamento de cromossomos eucarióticos por

recombinação, 109-149

- com marcadores moleculares, 120

- - polimorfismos de comprimento de sequências

simples, 125

- - polimorfismos de nucleotídeo único, 122

- - - usando haplótipos de, 123

- - - cálculo de *crossings* múltiplos não-vistos, 132

- de centrômero com tétrades lineares, 127

- diagnóstico de ligação, 111

- - como os *crossings* produzem recombinantes

para genes ligados, 112

- - evidência de que o *crossing over*, 113

- - - é um processo de quebra e reunião, 113

- - - ocorre no estágio de quatro cromátides, 113

- - múltiplos *crossings* podem incluir mais de

duas cromátides, 114

- - simbolismo de ligação e terminologia, 113

- - uso de frequência de recombinantes para

reconhecer ligação, 111

- por frequência, 114

- - cruzamento-teste de três pontos, 117

- - dedução da ordem dos genes por inspeção, 119

- - interferência, 119

- - unidades de mapa, 115

- - uso de proporções como diagnóstico, 120

- uso de mapas baseados em recombinação em

conjunto com mapas físicos, 134

- uso de valores Lod para avaliar a ligação em

heredogramas humanos, 131

- uso do teste do qui-quadrado para testar análise

de ligação, 129

Maquinaria transcricional, 340

Marcadores

- fenotípicos e moleculares mapeados no

cromossomo humano 1, 128

- microssatélites, 125

- minissatélites, 125

- moleculares, mapeamento com, 120

- - polimorfismos de comprimento de sequências

simples, 125

- - polimorfismos de nucleotídeo único, 122

- - - usando haplótipos de, 123

Marcas epigenéticas, 344, 351

Material genético, 226

- do fago, 229

- e DNA, 226

- - descoberta da transformação, 226

- - experimento de Hershey-Chase, 227

Matrizes de leitura abertas, 396

McClintock, experimentos de, 418

Meiócitos, 35

Meiose, 35

- complexos sinaptonêmicos na, 39

- *crossing* durante a, 463

- distribuição independente de cromossomos

na, 87

- DNA e transmissão gênica na mitose e, em

eucariontes, 45

- estágios da, 42

- - e mitose, 38

- linear de *Neurospora*, 88

Melanina, 9

Melanismo, 592

Melanócitos, 9

Mendel

- cruzamentos de, programa de, que produziu

uma proporção 9:3:3:1, 77

- lei de, 30, 76

Menino da bolha, doença do, 417

Merozigoto, 162

Meselson-Stahl, experimento de, 235

Metabolismo, 195

- da lactose, 305

- erros hereditários do, descobridor dos, 195

Método(s) (*v.* *th.* Técnica)

- ChIP, 407

- de introdução de um transgene, 631

- de sequenciamento didesoxi, 625

Micrografias eletrônicas de células normais e

transformadas pelo vírus do sarcoma de

Rous, 467

Microrganismos, trabalhando com, 153

MicroRNA, 257, 379

Milho, 628

- cromossomo 9 de, 418

- descoberta dos elementos de transposição

em, 418

- - autônomos e não-autônomos, 421

- - experimentos de McClintock, 418

- domesticado, principais diferenças morfológicas

entre teosinto selvagem e, 629

- grãos de, fenótipos produzidos por elementos

de transposição em, 421

- identificando um gene importante para a

domesticação do, 628

- planta de, semi-estéril, pólen de uma, 500

- variação genética na cor dos grãos de, 1

- vigor híbrido em, 85

miRNA (*v.* MicroRNA)

Mitocôndria(s), 16
 - células mostrando nucleóides dentro de, 94
 - mapas de DNA para, e cloroplastos, 94
 Mitose, 40
 - ação dos microtúbulos na, 48
 - DNA e transmissão gênica na, e meiose em eucariontes, 45
 - estágios da, 40
 - - meiose e, 38
 - pós-meiótica, 86
 Modificação genética, técnicas de, 649
 - *Arabidopsis thaliana*, 655
 - *Caenorhabditis elegans*, 657
 - *Drosophila melanogaster*, 659
 - *Escherichia coli*, 649
 - *Mus musculus*, 661
 Molécula(s), 610
 - de DNA, 44
 - - detecção de, 12
 - - modelo de Watson-Crick da, 226
 - - recombinante, 610
 - - - formação de uma, 612
 - de RNA, 287
 Moléculas-filhas, 242
 Moluscos, 581
 Monômero de ubiquitina, 461
 Monometil lisina, 349
 Monoplóides, 477, 481
 Morcego, estruturas ósseas de uma asa de, uma asa de pássaro e um braço e mão humanos, 599
 Mosaicismo, 422
 Mosca-da-fruta, 593
 - expressão gênica de um desenvolvimento de embrião da, 355
 - manchas nas asas da, 593
 mRNA (v. RNA mensageiro)
 mtDNA (v. DNA mitocondrial)
 Mudança tautomérica, 445
Mus musculus, 15, 660
 - ciclo de vida, 660
 - transgênese em, 635
 Mutação(ões), 503
 - auxotrófica, isolamento de uma, para arginina, 196
 - citoplasmáticas em humanos, 97
 - cromossômicas, 476
 - - identificação das, pela genômica, 503
 - - índice geral de, humanas, 503
 - - tipos de, 477
 - de pelagem amarela em labrador, epistasia recessiva devida a, 205
 - de quebra bifilamentar, 461
 - de sentido trocado, 442
 - espontâneas, 443
 - genéticas constitutivas, 307
 - gênica(s), 320
 - - e evolução molecular, 588
 - - haplossuficientes, 189
 - - polares, 311
 - - *vermillion*, 9
 - indel, 446
 - induzidas, 443
 - oncogênicas, 467
 - que interagem e afetam a cor do olho em *Drosophila*, 10
 - sem sentido, 442
 - sinônimas, 442
 - sítios, 47
 - - gênicos sensíveis a, 47
 - - no mtDNA em algumas doenças humanas, 98
 - vazantes, 44
 Mutação, reparo e recombinação, 439-473
 - base molecular das mutações, 443
 - - espontâneas, 443
 - - - em humanos, doenças de repetições de trinucleotídeos, 448
 - - - mecanismos das mutações espontâneas, 445
 - - - teste de flutuação de Luria-Delbrück, 443

- - induzidas, 449
 - - - mecanismos de mutagenese, 449
 - - - teste de Ames, 451
 - câncer, uma consequência fenotípica importante da mutação, 467
 - consequências fenotípicas, 440
 - - das mutações no DNA, 440
 - - - moleculares das mutações de ponto em uma região codificante, 442
 - - - moleculares das mutações de ponto em uma região não-codificante, 442
 - - - tipos de mutação de ponto, 440
 - - tipos de mutação de ponto, 440
 - mecanismo de *crossing over* meiótico, 463
 - - análise genética das tétrades e indícios do mecanismo de recombinação, 464
 - - modelo de quebra bifilamentar para recombinação meiótica, 465
 - - quebras bifilamentares programadas e a recombinação meiótica, 463
 - mecanismos biológicos de reparo, 453
 - - de quebras bifilamentares, 461
 - - por excisão, 454
 - - - de base, 454
 - - - de nucleotídeo, 456
 - - pós-replicação, 458
 - - propenso a erro, 459
 - - reversão direta de DNA danificado, 454
 Mutagenese, 408
 - aleatória, genética reversa e, 408
 - genética reversa e, direcionada, 409
 - mecanismos de, 449
 - técnica de, insercional, 175
 Mutágenos, 443
 - avaliação de, em nosso ambiente, 451
 Mutante(s), 154, 356
 - *Antennapedia*, 256
 - auxotróficos, 154
 - de genes de segmentação, 367
 - duplos, análise de, por mutações aleatórias, 201
 - homeóticos de *Drosophila melanogaster*, 356
 - mitocondrial de fenótipo poky, herança materna de, 95
 - operadores diplóides heterozigotos, síntese de β -galactosidase e permease em haplóides e, 308

N

Nematódeo *Caenorhabditis elegans*, 498
Neurospora, 653
 - *crassa*, 26, 89, 652
 - - ciclo de vida de, 88, 652
 - - fungo, 89
 - - via de síntese do pigmento carotenóide laranja em, 653
 - estágios de um cruzamento de, 88
 - meiose linear de, 88
 - segregação aberrante em, 465
 - tipo selvagem, 653
 - - e mutante crescendo em uma placa de Petri, 653
 - vias biossintéticas em, 195
Nicotiana longiflora, 546
 - linhagens de, 547
 Nocaute(s) gênico(s), 636, 651
 - *Arabidopsis thaliana*, 655
 - *Caenorhabditis elegans*, 657
 - células produtoras contendo um, direcionado, 637
 - direcionados, 649
 - *Drosophila melanogaster*, 659
 - *Escherichia coli*, 649
 - *Mus musculus*, 661
 - *Neurospora crassa*, 653
 - produção de camundongo contendo o, direcionado, 638
 - *Saccharomyces cerevisiae*, 651
 Nomenclatura genética, 663
 Northern, transferência de, 13
 Nucléolo, 28
 Nucleossomo, 28, 339
 - estrutura do, 339
 - montagem de, durante a replicação do DNA, 243
 Nucleotídeo(s), 229
 - do DNA, estruturas dos quatro, 230
 - percentagem de, que diferem entre humanos e chimpanzés, 11
 - pirimidínicos, 230
 - reparo, 456
 - - genômico por excisão de, acoplado à transcrição, 456
 - único, polimorfismos de, 122
 Número variável de repetições em tandem, 523

O

Óctade, 127
 Octâmero de histonas, 338
 Okazaki, fragmentos de, 238, 245
 Oligonucleotídeo, 406
 Oncogene(s), 467
 - células normais e células transformadas por um, 467
 - classes de, 468
 - *ras*, 468
 - *src*, 467
 Oncoproteína, 468
 Óperon(s), 305
 - *ara*, mapa do, 315
 - complexos, 317
 - de arabinose, 313
 - de triptofano, 314
 - *lac*, 305
 - - controle catabólico do, 312
 - - região controladora do, 314
 - - regulação do, 306
 - - repressão catabólica do, controle positivo, 311
 - peptídeos-líder de, de biossíntese de aminoácidos, 318
 - *trp*, 315
 - - modelo de atenuação do, 317
 - - ordem do gene no, 316
 - - transcrição do, 315, 317
 Orelhas com pêlos, 62
 Organela(s), 16, 94
 - celulares, 597
 - eucarióticas, 16
 - geneticamente distintas, 97
 - genomas de, 94
 Organismo(s)
 - diplóides, 85
 - eucarióticos, replicação do DNA em, 241
 - experimentais, método padrão para testar herdabilidade em, 558
 - geneticamente modificados, 632
 - haplóides, 86
 - multicelulares, 14
 - não-modelo, inserção de transgenes em, 411
 - transgênico, 631
 Organismo(s)-modelo, 13, 647
 - *Arabidopsis thaliana*, 14, 26, 654
 - *Caenorhabditis elegans*, 15, 378, 656
 - camundongo, 194
 - *Drosophila*, 358
 - - *melanogaster*, 14, 52, 658
 - *Escherichia coli*, 14, 155, 648
 - leveduras, 333
 - milho, 419
 - *Mus musculus*, 660
 - *Neurospora crassa*, 26, 89, 652
 - *Saccharomyces cerevisiae*, 650
 Ovelha Dolly (v. Dolly)

P

Palíndromo de DNA, 611
 Pão, levedura do, ciclo de vida, 333
 Pareamentos códon-anticódon permitidos pelas regras da oscilação, 285
 Parentes, similaridade fenotípica entre, 557
 Partenogênese, 477
 Pássaro, estruturas ósseas de uma asa de morcego, uma asa de, e um braço e mão humanos, 599
 Pedras de lava, 591
 Peixe *Astyanu mexicanus*, 593
 Pelagem amarela de labrador, mutação de, epistasia recessiva devida a, 205
 Pêlos, orelhas com, 62
 Peptídeos-líder de óperons de biossíntese de aminoácidos, 318
 Perkins, fórmula de, 133
 Pesquisa(s), 15
 - BLAST, 398
 - genética, 15
 - genômicas, 397
 Petri, placa de, *Neurospora* tipo selvagem e mutante crescendo em uma, 653
 Petúnia, flores de, demonstrando co-supressão, 269
 Piebaldismo, manchas de, dominante, 59
 Pigmento(s), 9
 - carotenóide laranja em *Neurospora*, via de síntese do, 653
 - herdados e sintetizados independentemente, 203
 - melanina, 9
 - *vermilion*, 9
 Pirimidina(s), 229, 452
 - pareamento de purinas com, 233
Pisum sativum, 30
 Placa de Petri, *Neurospora* tipo selvagem e mutante crescendo em uma, 653
 Planta(s), 191, 553
 - alopóliplóides, 481
 - campainha, flores da, 200
 - cruzamentos usando flores de uma, variegada, 96
 - de milho semi-estéril, pólen de uma, 500
 - dicotiledôneas, 596
 - distribuição fenotípicas de genótipos de três, 550
 - engenharia genética em, 631
 - homozigota, 34
 - masculina e feminina, 51
 - monoplóides, 484
 - transgênica, geração de uma, 634
 Plaqueamento réplica, técnica de, 444
 Plasmídeo(s), 158
 - com segmentos de muitos hospedeiros bacterianos anteriores, 166
 - construído como vetor de clonagem de DNA, 649
 - determinantes genéticos vindos de, 165
 - F que levam fragmentos genômicos, 163
 - inserção de um gene em um, de DNA recombinante, 612
 - integrativo de levedura, 651
 - R, 164, 423
 - - com genes de resistência levados em um transposon, 166
 - sistema de, Ti, 632
 - Ti, infecção pelo, 633
 - vetor, 614
 Poisson, distribuição de, 132
 Pólen, 500
 - de uma planta de milho semi-estéril, 500
 - normal e abortado de uma translocação heterozigota, 500
 Polidactilia, 58, 382
 Poligenes, 91
 - histograma de, de uma autofecundação, 93
 - - diíbrida, 93
 - - triíbrida, 93
 - na prole de uma autofecundação diíbrida, 92
 Polimerase(s), 460

- reação em cadeia da, 622
 - translesão, 460
 Polimorfismo(s), 58
 - autossômicos, 58
 - balanceado, 537
 - de comprimento, 123
 - - de fragmento de restrição, 123
 - - de seqüências simples, 125
 - de cor de conchas em *Liguus fascitus*, 517
 - de nucleotídeo único, 122
 - - usando haplótipos de, 123
 - de proteínas, 520
 - genético, 8
 Polipeptídeo, 7, 197, 277
 - uso de, e similaridade de DNA, 396
 Poliplóides, 477
 Poliploidia, 479, 595
 - ancestral, duplicações no genoma de *Saccharomyces* por, 495
 População(ões)
 - alopátricas, 603
 - genética de, 517-544
 - - efeitos da reprodução sexual na variação, 525
 - - - endogamia e reprodução preferencial, 528
 - - - heterozigose, 527
 - - - reprodução aleatória, 527
 - - - segregação meiótica e equilíbrio genético, 525
 - variação genética, 581
 - dentro de, e entre populações, 588
 - e divergência de, 581
 - e sua modulação, 518
 - - - estrutura do DNA e polimorfismo de seqüências, 521
 - - - observações da variação, 518
 - - - polimorfismos de proteínas, 520
 Primase, 238
 Primer, 238
 Primossomo, 238
 Procariontes, 14, 152, 258, 289
 - elementos de transposição em, 422
 - início da tradução em, 289
 - início da transcrição em, 260
 Produtos, 443
 - aneuplóides, não-disjunção meiótica para, 485
 - gênicos, conseqüências das mutações de ponto nos, 443
 - meióticos de um trissômico, 486
 Profago, 171
 Programa da investigação genética, 7
 - começando, 8
 - - com o DNA, genética reversa, 11
 - - com variação, genética direta, 8
 - necessidade da variação, 7
 Projeto Genoma, 389
 Prole, 481
 - de um cruzamento de couve e rabanete, 481
 - síndrome de Down na, de uma translocação heterozigota, 501
 Proliferação celular, antígeno nuclear de, 242
 Promotores gênicos, 332, 337
 Proteína(s), 4
 - acessórias, 239
 - ativadora(s), 312
 - - do catabolismo, 312
 - - transcricionais, 334
 - *Bicoid*, 370
 - CAF-1, 243
 - de ligação, 263
 - - unifilamentares, 241
 - dobramento da, dentro da célula, 292
 - *Duffy*, 595
 - e a diferença na taxa de mutação, 589
 - e sua síntese, 275-300
 - - código genético, 279
 - - colinearidade de gene e proteína, 279
 - - estrutura, 277
 - - proteoma, 291
 - - ribossomos, 286

- - tRNA, o adaptador, 283
 - especificando a seqüência de aminoácidos de uma, 5
 - estudo de interações de, com o uso do sistema diíbrido de levedura, 407
 - eucarióticas, distribuição de, de acordo com amplas categorias de funções biológicas, 600
 - fibrosas, 279
 - fosforilação e desfosforilação de, 294
 - funcionando na forquilha de replicação, 240
 - Gal, 334, 336
 - *gap*, 372
 - globulares, 279
 - *Hox*, 363
 - - de *Drosophila* e a de vertebrados e similaridades, 363
 - interação de, em um organismo, 294
 - isoformas de, 292
 - MC1R, 592
 - MutH, 458
 - NHEJ, 462
 - níveis de estrutura da, 278
 - PCNA, 460
 - polimorfismos de, 520
 - reguladora(s), 204
 - - combinações de, e controle dos tipos celulares, 343
 - - interação de, e seu alvo, 204
 - regulatórias, 304
 - - ligação específica de seqüência de, ao DNA, 321
 - relação entre genética e mudança funcional na evolução de, 590
 - - inativação gênica, 592
 - - morfológica, 590
 - - sinal característico da seleção positiva nas seqüências de DNA, 590
 - remodeladoras de cromatina e ativação gênica, 338
 - repórter verde fluorescente transgênica, expressa nos olhos de alguns insetos não-modelo, 411
 - repressora, 305
 - separação eletroforética revelando variantes de, 520
 - seqüência de genes e, 280
 - sondagem de uma, específica, 13
 - tipo selvagem, funções de, e propriedades de mutações promotoras de tumor nos genes correspondentes, 469
 Proteína-1 de heterocromatina, 349
 Proteínas-DNA, 342
 Proteoma(s), 265, 291, 395
 - comparando os, entre espécies distantes, 600
 Protocolo experimental para localizar gene, 567
 Proto-oncogenes, 468
 Provírus, 426
 Pseudo-acondroplasia humana, fenótipo de, 57
 Pseudodominância, mapeamento de alelos mutantes por, 493
 Pseudogenes, 399, 597
 - processados, 399
 Punção de vilosidades coriônicas, 630
 Punnett, quadrado de, 78
 Purinas, pareamento de, com pirimidinas, 233

Q

QI, 560
 - herdabilidade do, 557
 - teste de, 560
 Quadrado de Punnett, 78
 Quebra bifilamentar, mutação de, 461
 Quiasma(s), 463
 - e locais de *crossing*, 112
 Qui-quadrado, teste de, uso do, 81
 - nas proporções monoíbridas e diíbridas, 81
 - para testar análise de ligação, 129

R

Radiação ionizante, 451
 Radicais oxigênio, 448
Raphanobrassica anfidiplóide, origem de, 481
 Reação em cadeia da polimerase, 622
 Rearranjos cromossômicos, origens dos, 491
 Recombinação
 - homóloga não-alélica, 490
 - reparo e mutação (*v.* Mutação, reparo e recombinação)
 Regras de Chargaff, 229
 Regulação gênica, 303 (*v. tb.* Expressão gênica, regulação da)
 - teste do papel de um elemento conservado na, 403
 Relógio molecular, 589
 Reparo genômico, 454
 - de mal pareamento, 458
 - - em *Escherichia coli*, 459
 - global, 456
 - sistema de, 454
 - - dependentes de homologia, 454
 - - por excisão, 454
 - - - de base, 454
 - - - de nucleotídeo acoplado à transcrição, 456
 Replissomo, 239
 - eucariótico, 242
 Repressão catabólica do óperon *lac*, 311
 Reprodução sexual, 525
 - efeitos da, na variação genética, 525
 - - endogamia e reprodução preferencial, 528
 - - heterozigose, 527
 - - reprodução aleatória, 527
 - - segregação meiótica e equilíbrio genético, 525
 Retrotransposons, 425
 - e características em comum com retrovírus, 427
 - LTR, 428
 Retrovírus, 426
 - ciclo de vida, 427
 - retrotransposons e características em comum com, 427
 Revestimento telomérico, estrutura do, 247
 Revolução genômica, 388
 Ribonucleotídeos, 256
 Ribose, 255
 Ribossomo(s), 6, 286
 - sítios principais de interação no, 288
 Ribozima, 255
 RNA, 12, 253-273
 - bifilamentar, um modo para gerar, em uma célula, 270
 - de interferência, 257, 269
 - - perturbação da função gênica com o uso de, 409
 - funcionais, 256, 266
 - - auto-recomposição de íntrons e o mundo de, 267
 - - pequenos RNA, 266
 - - - de interferência, 268
 - - - nucleares, mecanismo de recomposição dos éxons, 266
 - intermediário, 427
 - - demonstração de transposição por um, 428
 - mensageiro, 6, 256
 - - padrões complexos de recomposição eucariótica de, 265
 - - regulação da tradução do, e linhagem celular em *Caenorhabditis elegans*, 377
 - microRNA 257
 - moléculas de, 287
 - mundo de, 268
 - nucleares, 256, 266
 - polimerase, 303, 331
 - primeiros experimentos e o, intermediário, 254
 - processamento co-transcricional do, 264
 - propriedades do, 255
 - quatro ribonucleotídeos encontrados no, 256
 - recomposição do, e determinação do sexo em *Drosophila*, 376

- ribossômicos, 256, 276
 - seqüências de DNA e, transcrito, 259
 - sítios de ligação de, 397
 - sondagem de um, específico, 13
 - sondas de, 618
 - transcrição, 257
 - - em eucariontes, 261
 - - - alongamento, término e processamento pré-m-, 263
 - - - início, 262
 - - estágios da, 258
 - - - alongamento, 260
 - - - iniciação em procariontes, 258
 - - - término, 261
 - - visão geral, 257
 - transportador, 6, 256, 276, 283
 - - estrutura do, 284
 - - que podem servir de códons de serina, 286
 Rosalind Franklin, experimento de, 231
 Rous, sarcoma de, 467
 rRNA (*v.* RNA ribossômico)

S

Saccharomyces, 495
 - *cerevisiae*, 333, 495, 650
 - - ciclo de vida, 650
 - - engenharia genética em, 631
 - duplicações no genoma de, por poliploidia ancestral, 495
Salmonella typhimurium, 586
 Sanger, seqüenciamento de, 624
 Sarcoma de Rous, 467
 Segregação, 38, 525
 - citoplasmática, 95
 - - modelo de, 97
 - cromossômica, demonstração da, ao nível molecular, 38
 - meiótica e equilíbrio genético, 525
 Seleção, 564
 - familiar, 564
 - purificadora, 589
 Sementes de *Crinum longifolium*, 569
 Semi-esterilidade, 499
 Seqüência(s), 288
 - *contig*, 392
 - de Sanger, 624
 - de Shine-Dalgarno, 288
 - didesoxi, método de, 625
 Seqüência genômica, 391
 - de clone ordenado, 391
 - deduzindo os genes codificantes de proteína, 395
 - mapa de, criando um, 389
 - - estabelecendo uma biblioteca genômica de clones, 391
 - - preenchendo espaços na seqüência, 395
 - - transformando leituras de seqüências em um mapa de seqüência, 389
 - - usando o enfoque de clone ordenado para seqüenciar um genoma complexo, 394
 - - usando o enfoque *shotgun* de genoma inteiro, 391
 - - - para criar uma seqüência-rascunho de um genoma complexo, 392
 Serina, diferentes RNA transportadores que podem servir de códon de, 286
 Sexo, 51
 - heterogamético, 51
 - homogamético, 51
 Shine-Dalgarno, seqüências de, 288
 Silenciamento genético, 346
 - causado pela expansão da heterocromatina, 348
 - de um cromossomo inteiro, 350
 - mudança do tipo reprodutivo e, 346
 - transgênico, 269
 Símbolos genotípicos usados em genética de bactérias, 154

Similaridade fenotípica entre parentes, 557
 Síndrome(s)
 - de Cockayne, 456
 - de Down, 476, 486
 - - características da, 487
 - - efeito da idade materna na, 487
 - - na prole de uma translocação heterozigota, 501
 - de feminização testicular, 62
 - de Klinefelter, 486
 - - características da, 487
 - de Turner, 485
 - de Werner, 247
 - de Williams, 493
 - - origem de deleção da, 494
 - do *cri du chat*, 493
 - - origem de deleção da, 494
 - do X frágil, 448
 - HNPCC, 459
 Sintenia, mapas de, do genoma humano, 601
 siRNA (*v.* RNA de interferência)
 Sistema(s)
 - biologia de, 408
 - de plasmídeo Ti, 632
 - de reparo genômico, 454
 - - dependentes de homologia, 454
 - - por excisão de base, 454
 - diíbrido de leveduras, estudo de interações de proteína com o uso de, 407
 - GAL, 334
 - *lac*, descoberta do, controle negativo, 306
 - - análise genética do promotor *lac*, 309
 - - caracterização molecular do repressor *lac* e operador *lac*, 310
 - - evidência genética, 307
 - - - da alosteria, 309
 - - - do operador e repressor, 307
 - - genes controlados juntos, 307
 - - mutações polares, 311
 - MNS, 527
 - SOS, 460
 Sítio(s)
 - alostérico, 304
 - apurínicos, 447
 - de mutações no mtDNA em algumas doenças humanas, 98
 - gênicos sensíveis a mutação, 47
 SNP (*v.* Polimorfismos de nucleotídeo único)
 snRNA (*v.* RNA nucleares)
 Sondagem, 13
 - de um RNA específico, 13
 - de uma proteína específica, 13
 Sondas, 12, 618
 - de DNA, 618
 - de RNA, 618
 - para um DNA específico, 12
 Southern, transferência de, 13, 620
 Spliceossomo, 254
 - montagem e funcionamento do, 267
 SpoII, enzima, 463
Streptococcus pneumoniae, 226
 Substratos hemimetilados, 351
 Subunidades polipeptídicas, 278
Supercontigs, 393
 Supressores, 206

T

Tandem, número variável de repetições em, 523
 Tautômeros, 445
 Técnica(s) (*v. tb.* Método)
 - de DNA recombinante, 620
 - de manipulação genética, 651
 - - *Neurospora crassa*, 653
 - - *Saccharomyces cerevisiae*, 651
 - de modificação genética, 649
 - - *Arabidopsis thaliana*, 655
 - - *Caenorhabditis elegans*, 657

- - *Drosophila melanogaster*, 659
- - *Escherichia coli*, 649
- - *Mus musculus*, 661
- de mutagênese insercional, 175
- de plaqueamento réplica, 444
- Telomerase, 245
- Telômeros, 28
- câncer e envelhecimento, 246
- e telomerase, 245
- Tentilhões, espécies de, 582
- Teosinto selvagem, principais diferenças morfológicas entre, e milho domesticado, 629
- Terapia gênica, 417
- de linhagem germinativa, 640
- em camundongo, 639
- em mamíferos, 640
- humana, 639
- somática, 640
- Teste(s)
- de Ames, 451
- de complementação, 199
- - definição de conjuntos de genes usando o, 199
- - usando um heterocáron, 203
- de flutuação de Luria-Delbrück, 443
- de QI, 560
- diíbrido para estudar o interatoma, 405
- do papel de um elemento conservado na regulação gênica, 403
- do qui-quadrado, uso do, 81
- - nas proporções monoíbridas e diíbridas, 81
- - para testar análise de ligação, 129
- Tétrade(s), 36
- análise genética das, e indícios do mecanismo de recombinação, 464
- Thermus aquaticus*, 622
- Timina, 4, 451
- Tirosina, 9
- Tomate(s), 84
- cruzamento de, 84
- mapa dos 12 cromossomos de, 121
- Topoisomerases, 241
- Transcrição alostérica, 305
- Transcriptase reversa, 426
- Transcriptoma, estudo do, uso de microarranjos de DNA para o, 405
- Transdução, processo de, 171
- descoberta da, 171
- especializada, 173
- - mecanismo de, 174
- generalizada, 171
- Transferência, 13
- de Northern, 620
- de Southern, 13, 620
- de Western, 13
- Transformação bacteriana, 167
- Transgene(s), 411
- inserção de, em organismo não-modelo, 411
- métodos de introdução de um, 631
- Transgênese, 634
- *Arabidopsis thaliana*, 655
- *Caenorhabditis elegans*, 634, 657
- *Drosophila melanogaster*, 634, 659
- *Escherichia coli*, 649
- *Mus musculus*, 635, 661
- *Neurospora crassa*, 653
- *Saccharomyces cerevisiae*, 651
- Translocação(ões), 500
- cromossômicas, 499
- - recíprocas, 499

- - robertsonianas, 499
- heterozigota, 500
- - pólen normal e abortado de uma, 500
- - pseudoligação de genes em uma, 501
- somática, câncer por, 503
- Transmissão gênica, DNA e, na mitose e meiose em eucariontes, 45
- Transposase, 635
- Transposon(s), 431
- características estruturais de, compostos e simples, 423
- de DNA, 428
- - utilidade dos, para a descoberta do gene, 431
- plasmídeo R com genes de resistência levados em um, 166
- procarióticos, 423
- Tribolium castaneum*, 411
- Trigo do pão, origem proposta para o, por anfídiploidia ancestral, 483
- Trimetil lisina, 349
- Trinucleotídeos, doenças de repetição de, 448
- Triplóide, pareamento meiótico em, 479
- Triptofano, 314
- códons de, 317
- óperon de, 314
- Trissômicos, 486
- de *Datura*, 488
- produtos meióticos de um, 486
- Triticale*, 483
- Triturus viridescens*, 258
- tRNA (*v* RNA transportador)
- Turner, síndrome de, 485

U

- Ubiquitina, 294
- monômero de, 461
- Ubiquitinização, 294
- Uracil, 255
- Uvas, 484
- diplóides, 484
- tetraplóides, 484

V

- Valores Lod, uso de, para avaliar a ligação em heredogramas humanos, 131
- Variação genética, 8, 586
- dentro de populações e entre populações, 588
- e divergência de populações, 581
- e sua modulação, 518
- - estrutura do DNA e polimorfismo de seqüências, 521
- - observações da variação, 518
- - polimorfismos de proteínas, 520
- efeitos da reprodução sexual na, 525
- - endogamia e reprodução preferencial, 528
- - heterozigose, 527
- - reprodução aleatória, 527
- - segregação meiótica e equilíbrio genético, 525
- fontes de, 530
- - por migração, 532
- - por mutação, 530
- - por recombinação, 531
- herdabilidade da, 586
- na cor dos grãos de milho, 1
- Variança, 558
- ambiental, 558

- genética, 558, 562
- - estimando os componentes da, 562
- Vertebrados, 363
- número de substituições de aminoácidos na evolução dos, 589
- proteína *Hox* de *Drosophila* e de, e similaridades, 363
- Vetores, 392
- BAC, 614
- bacteriófagos, 614
- de clonagem, escolha dos, 613
- de leveduras simples, 632
- PAC, 614
- para insertos maiores de DNA, 614
- plasmidiais, 613
- YAC, 614
- Vias, interação gênica em, 195
- biossintéticas em *Neurospora*, 195
- tipos de vias, 197
- Vigor híbrido, 84
- em milho, 85
- Vilosidades coriônicas, punção de, 630
- Vírus, 14, 152
- do sarcoma de Rous, micrografias eletrônicas de varredura de células normais e células transformadas pelo, 467
- genética de bactérias e seus, 151-186
- - conjugação bacteriana, 154
- - - descoberta da conjugação, 155
- - - descoberta do fator fertilidade, 156
- - - linhagens Hfr, 157
- - - mapeamento de cromossomos bacterianos, 161
- - - plasmídeos F que levam fragmentos genômicos, 163
- - - plasmídeos R, 164
- - genética de bacteriófagos, 167
- - - infecção de bactérias por fagos, 168
- - - mapeamento de cromossomos de fagos usando cruzamentos de fagos, 169
- - mapas físicos e mapas de ligação comparados, 174
- - trabalhando com microrganismos, 153
- - transdução, 171
- - - descoberta da, 171
- - - especializada, 173
- - - generalizada, 171
- - transformação bacteriana, 167

W

- Watson-Crick, 226
- estrutura do DNA antes de, 229
- modelo de, da molécula de DNA, 226
- Werner, síndrome de, 247
- Western, transferência de, 13
- Williams, síndrome de, 493
- origem de deleção da, 494

X

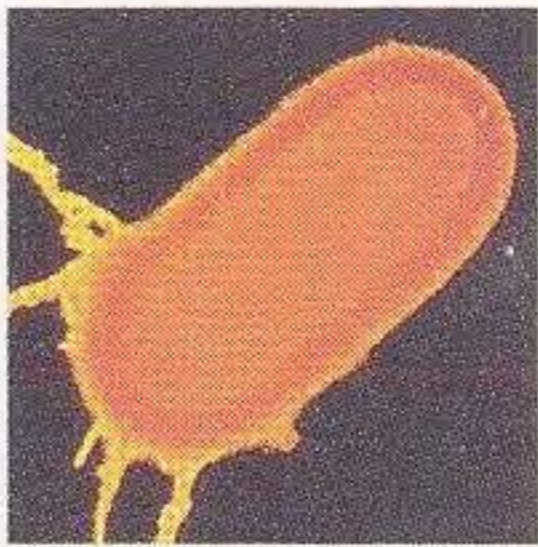
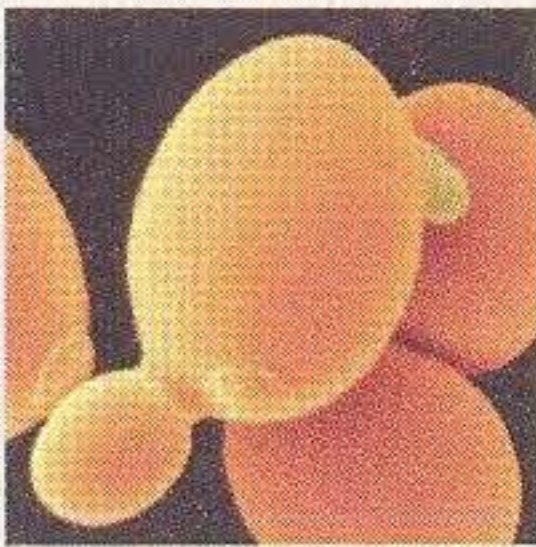



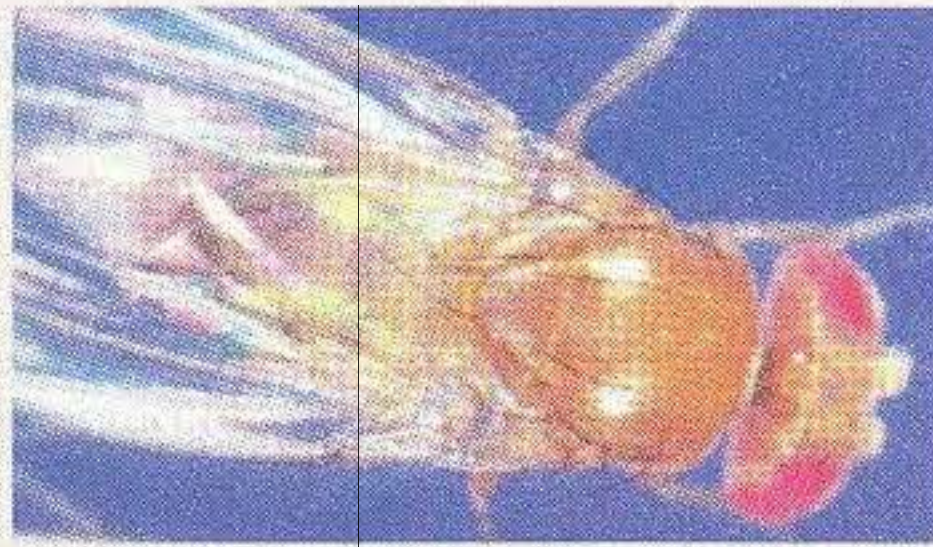

- Xantina desidrogenase, 520
- Xeroderma pigmentoso, 440, 456

Z

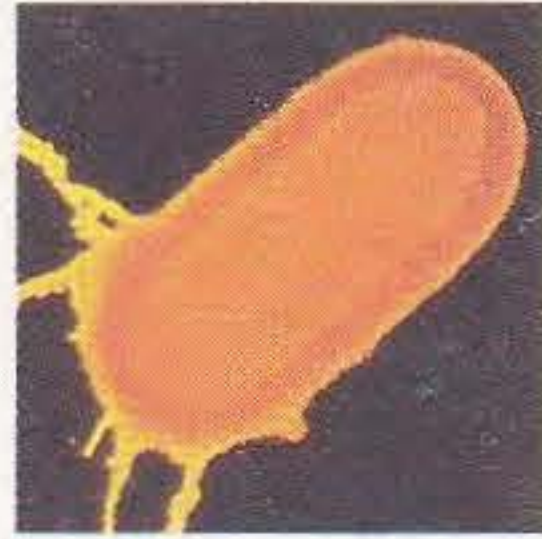
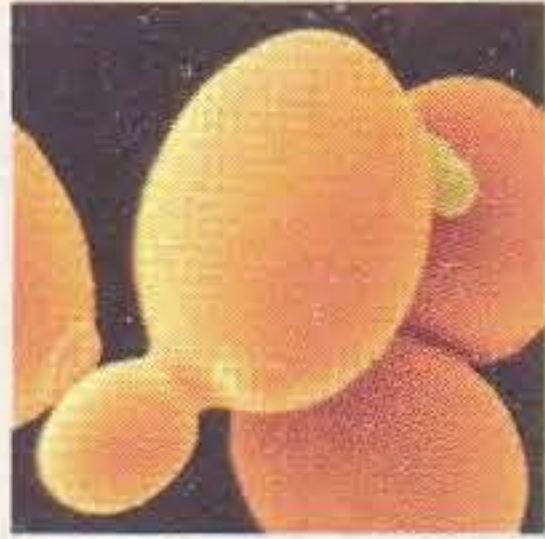


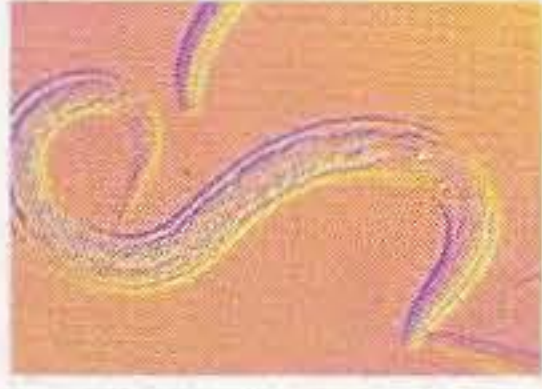


- Zigoto, 34

Índice de Organismos-modelo

O seguinte quadro fornece referências das páginas para discussões de organismos-modelo no texto.

							
	Bactéria (<i>E. coli</i>)	Levedura do pão (<i>S. cerevisiae</i>)	Mofo do pão (<i>N. crassa</i>)	Mostarda (<i>A. thaliana</i>)	Nematelminto (<i>C. elegans</i>)	Mosca-das-frutas (<i>D. melanogaster</i>)	Camundongo (<i>M. musculus</i>)
PRINCIPAIS CARACTERÍSTICAS pp. 647-661	p. 648	p. 650	p. 652	p. 654	p. 656	p. 658	p. 660
DESTAQUES	p. 155	p. 333	p. 89		p. 378	p. 52, pp. 358-359	p. 194
CAPÍTULOS							
1. O Enfoque Genético à Biologia	descrição do organismo, p. 14	descrição do organismo, p. 14	descrição do organismo, p. 14	descrição do organismo, p. 14	descrição do organismo, p. 15	análise genética da cor de olho, p. 10 descrição do organismo, p. 14 temperatura e desenvolvimento, p. 18 ruído de desenvolvimento, p. 19	descrição do organismo, p. 15
2. Herança Monogênica		número de genes, p. 30 análise genética usando asco, pp. 37-38	mutantes de desenvolvimento do micélio, p. 26 origem do DNA, p. 30	mutantes de desenvolvimento da flor, pp. 26, 48-49		identificação de um gene para desenvolvimento da asa, p. 48 determinação do sexo, pp. 50-51 herança ligada ao X (cor do olho), pp. 86-87	
3. Distribuição Independente de Genes			ciclo vital, p. 86 observação de distribuição independente, pp. 85-89 herança materna (mutantes poky), pp. 95-96 segregação citoplasmática, p. 95				
4. Mapeamento de Cromossomos Eucarióticos por Recombinação			mapeamento do centrômero, pp. 127-129 análise da tétrade dos crossings, p. 113			mapa do cromossomo de <i>Drosophila</i> , p. 103 experimentos de Morgan, ligação, pp. 111-112 análise de ligação diíbrida, pp. 116-117 cruzamento-teste de três pontos, pp. 117-119 interferência, p. 119 sem crossing nos machos, p. 119	correlacionando haplótipos e características, p. 125
5. A Genética de Bactérias e seus Vírus	conjugação, p. 155 cruzamentos de fagos, p. 169 mapeamento do gene rII, pp. 169-170 experimentos clássicos de transdução, pp. 171-173 mapa genômico, pp. 174-175						

(Continua)

							
	(<i>E. coli</i>)	(<i>S. cerevisiae</i>)	(<i>N. crassa</i>)	(<i>A. thaliana</i>)	(<i>C. elegans</i>)	(<i>D. melanogaster</i>)	(<i>M. musculus</i>)
6. Interação Gênica		sinal de transdução para mudança do tipo reprodutivo, p. 198 supressão, p. 206 mutações modificadoras, p. 207	experimentos de Beadle-Tatum, pp. 195-197 complementação, p. 199			supressão de cor de olho, pp. 206-207	papel na determinação de interações gênicas humanas, p. 188 exemplo de gene de haploinsuficiência, p. 189 coloração de pelagem, pp. 192-193
7. DNA: Estrutura e Replicação	experimento de Hershey-Chase, p. 227 experimento de Meselson-Stahl, pp. 235-238 DNA polimerases, pp. 236-238 velocidade da replicação, p. 239 replissomo, pp. 239-241 origem da replicação, p. 242	replissomo, p. 242 origens de replicação, p. 242 controle do ciclo celular, p. 244					experimento de Griffith, p. 226
8. RNA: Transcrição e Processamento	número de genes, p. 253 experimento de pulso e caça de Volkin-Astrachan, p. 254 estágios da transcrição, pp. 258-261 seqüências promotoras, p. 258 fatores sigma, p. 260 término, p. 261 densidade gênica, p. 261	número de genes, p. 253 complexo da origem do reconhecimento, p. 263 RNA polimerase II, p. 263 frequência de íntrons, p. 264			silenciamento gênico por RNAi, p. 268	número de genes, p. 253 densidade gênica, p. 261	
9. Proteínas e sua Síntese	Yanofsky e a colinearidade, p. 279 experimento do tamanho do códon de Crick, p. 281	abundância de transcrição de RNA, p. 276 elucidação do código genético de Nirenberg, p. 282		número de genes de cinase, p. 283	interatoma, p. 294		
10. Regulação da Expressão Gênica em Bactérias e seus Vírus	experimentos do óperon <i>lac</i> de Jacob-Monod, p. 307 óperon <i>trp</i> , p. 315						
11. Regulação da Expressão Gênica em Eucariontes		Sistema GAL, pp. 334-335 mutantes SWI-SNF, pp. 338-339 modificação da histona, p. 339 controle do tipo reprodutivo, p. 342 silenciamento gênico e mudança do tipo reprodutivo, pp. 345-346				silenciamento epigenético, p. 348 variação do efeito de posição, pp. 348-350 compensação de dose, p. 351	imprinting genômico, pp. 343-344
12. O Controle Genético do Desenvolvimento					como organismo-modelo, p. 357 destinos de linhagem celular, pp. 377-379 tempo de desenvolvimento, p. 379	mutantes homeóticos, p. 356 como organismos-modelo, p. 357 genes homeóticos, pp. 357-360 desenvolvimento inicial, pp. 365-374 determinação do sexo, pp. 376-377 múltiplos papéis do gene <i>hedgehog</i> , pp. 380-381	como organismo-modelo, p. 357 grupos de gene <i>Hox</i> , p. 360

13. Genomas e Genômica	genômica comparativa de patogênicos e não-patogênicos, pp. 403-405 mutagênese sistemática direcionada, p. 408	ano de seqüenciamento, p. 388 preenchendo espaços de seqüência, p. 395 teste diíbrido, p. 405			ano de seqüenciamento, p. 388 vetores para mapeamento, p. 392 repetições em, p. 392	ano de seqüenciamento, p. 388 método de seqüenciamento, p. 392 tendenciosidade de códon, p. 398	genômica comparativa humano-camundongo, pp. 400-401 uso na identificação dos elementos codificantes conservados, p. 402
14. O Genoma Dinâmico: Elementos de Transposição	elementos IS, pp. 425-428	elementos <i>Ty</i> , pp. 425-428 falta de transposons do DNA, p. 431 redirecionamento de transposon, p. 434				elementos tipo <i>copia</i> , p. 428 elementos <i>P</i> , pp. 428-431 disgenesia híbrida, p. 428 elemento <i>P</i> para marcação/transgênese, p. 431 marcação de transposon, p. 536	vírus de tumor mamário de camundongo, p. 434 atividade de retrotransposon, p. 433
15. Mutação, Reparo e Recombinação	teste de flutuação de Luria-Delbrück, p. 443 metiltransferase, p. 454 hotspots de mutação, p. 455 reparo de mal pareamento, p. 458 reparo SOS, p. 460	teste diíbrido, p. 462 recombinação meiótica, p. 463 número de quebras bifilamentares, pp. 464, 466	recombinação meiótica, p. 463				
16. Alterações Cromossômicas em Grande Escala		duplicação de todo o genoma, p. 495			cromossomos balanceadores, p. 498	poliplóides experimentais, p. 483 não-disjunção, p. 484 aneuploidia, p. 488 compensação de dose, p. 489 pseudodominância e mapeamento de deleção, p. 492 inversões, p. 497 cromossomos balanceadores, p. 498 variação de efeito de posição, p. 502	
17. Genética de Populações						variação do sítio de restrição, pp. 522-523 polimorfismo de pares de bases sinônimas, p. 524	
18. Genética Quantitativa						normas de reação, p. 551 teste de herdabilidade, p. 559 distinção de mutações genéticas entre fenótipos similares, p. 565	
19. Genética Evolutiva	experimento de seleção, p. 586 evolução da resistência a antibiótico, p. 586					seleção para mudança no tamanho da asa, p. 583 características canalizadas, p. 587 distinções de espécies, p. 590 evolução da regulação gênica e manchas nas asas, pp. 593-594 origem gênica dos elementos de transposição, p. 598 genética do isolamento de espécies, p. 603	genômica comparativa humano-camundongo, p. 601
20. Isolamento e Manipulação de Gene	enzimas de restrição, p. 611 usadas em clonagem, p. 612 clonagem com bacteriófago, vetores, p. 613 clonagem com plasmídeos <i>Ti</i> , p. 632	cromossomo artificial de levedura (YAC), p. 614 engenharia genética com vetores de levedura, p. 631			tamanho do genoma e clonagem, p. 616 transgênese em, p. 634	transgênese em, p. 634 nocaute gênico direcionado, p. 636 terapia gênica, p. 639	

Sobre a Capa

Variação nos padrões das asas da borboleta *Morpho didius*. O macho (em cima, à esquerda) e a fêmea (em cima, à direita) têm padrões de asas marcadamente diferentes, controlados por genes de determinação do sexo. Aqui, um raro evento genético durante o desenvolvimento produziu dois "ginandromorfos" (embaixo), borboletas nas quais parte do tecido da asa mostra um padrão mais típico do sexo oposto. Fotos por cortesia de Nipam Patel, HHMI e University of California, Berkeley.

Novidades Desta Edição

A nona edição do *Introdução à Genética* foi completamente revisada para destacar como a genética é praticada hoje.

Esta nova edição apresenta

Acentuada ênfase em:

- Aplicação da análise genética a uma ampla gama de processos biológicos
- Recentes aplicações humanas

Ampla cobertura sobre:

- Regulação gênica eucariótica
- Genômica comparativa
- Experimentos de ponta



Conteúdo online (elaborado pela W.H. Freeman and Co.) disponível, totalmente em inglês, em www.editoraguanabara.com.br/pub/livro/griffiths/griffiths.html

Estão disponíveis 43 animações em vídeo e testes práticos para ajudar os estudantes em suas revisões, proporcionando assim um melhor aprendizado dos processos e conceitos genéticos importantes.



GUANABARA  KOOGAN
www.editoraguanabara.com.br

ISBN 978-85-277-1497-6



9 788527 714976